

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 18:13:48 ; Search time 13687 Seconds  
(without alignments)  
11466.729 Million cell updates/sec

Title: US-10-045-072-1  
Perfect score: 3621  
Sequence: 1 tggggggggtgatctctg.....tgatcgctgcttctctaa 3621

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues 6940544

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sv.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rtd.\*
- 36: em.htg.man.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3621	100.0	3621	6	AR123974	AR123974 Sequence
2	3621	100.0	3621	6	AR213173	AR213173 Sequence
3	3621	100.0	3637	1	AF038548	AF038548 Coryneb
4	3621	100.0	332050	1	AF005276	AF005276 Coryneb
5	3621	100.0	349980	6	AX127145	AX127145 Sequence
6	3578	98.8	3728	1	CGPYC	Y09548 Corynebacte
7	3578	98.8	3728	6	A97276	A97276 Sequence 1
8	3578	98.8	3728	6	BD140686	BD140686 Method fo
9	3469	95.8	3657	1	AF503915	AF503915 Coryneb
10	3420	94.4	3420	6	AX120849	AX120849 Sequence
11	3420	94.4	3420	6	BD162966	BD162966 Novel pol
12	3398.8	93.9	3474	6	AX453601	AX453601 Sequence
13	3398.8	93.9	3474	6	AX453603	AX453603 Sequence
14	2282.8	63.0	308750	1	AF005216	AF005216 Coryneb
15	2279.6	63.0	4013	6	BD094153	BD094153 Genes for
16	2279.6	63.0	4851	1	AB083299	AB083299 Coryneb
17	1927.4	53.2	341553	1	EX248355	EX248355 Coryneb
18	1708	47.2	1719	6	AX064993	AX064993 Sequence
19	1395	38.5	1406	6	AX064995	AX064995 Sequence
20	1344.6	37.1	291000	1	SCO999105	AF939105 Streptomy
21	1328.2	36.7	4218	1	AF262949	AF262949 Mycobacte
22	1294.4	35.7	17992	1	AE007125	AE007125 Mycobacte
23	1294.4	35.7	348676	15	EX842581	EX842581 Mycobacte
24	1292.8	35.7	318050	1	EX248344	EX248344 Mycobacte
25	1249.6	34.5	31175	1	MTU00024	U00024 Mycobacteri
26	1249.6	34.5	31176	1	MSGY2	AD000009 Mycobacte
27	1249.6	34.5	32155	6	AR345347	AR345347 Sequence
28	1033	28.5	1083	6	AX064989	AX064989 Sequence
29	1033	28.5	1083	6	AX064991	AX064991 Sequence
30	907	25.0	939	6	AX064985	AX064985 Sequence
31	907	25.0	939	6	AX064987	AX064987 Sequence
32	899	24.8	302938	1	EX251412	EX251412 Tropherym
33	899	24.8	324227	1	AE016852	AE016852 Tropherym
34	741.4	20.5	303862	1	AE017215	AE017215 Geobacter
35	720	19.9	4010	5	AB086371	AB086371 Pagrus ma
36	714	19.7	3994	4	AY185595	AY185595 Bos tauru
37	707.8	19.5	3537	4	AY225510	AY225510 Sus scrof
38	705.4	19.5	34786	1	EX640438	EX640438 Bordetell
39	703.2	19.4	3516	8	AF097728	AF097728 Aspergill
40	701.2	19.4	3559	5	AF509529	AF509529 Gallus ga
41	699	19.3	349146	1	EX640424	EX640424 Bordetell
42	691.8	19.1	3953	9	HSU04641	U04641 Human pyruv
43	691.8	19.1	4049	9	BC011617	BC011617 Homo sapi
44	690.6	19.1	4067	10	MUSMPYR	L09192 Mus musculu
45	690.2	19.1	3960	9	HSU30891	U30891 Human pyruv

ALIGNMENTS

RESULT 1	AR123974	AR123974	3621 bp	DNA	linear	PAT 16-MAY-2001
LOCUS	Sequence 1	from patent US 6171833				
DEFINITION	AR123974					
ACCESSION	AR123974					
VERSION	AR123974.1	GI:14109335				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3621)					
AUTHORS	Sinskey,A.J., Lessard,P.A. and Willis,I.B.					
TITLE	Pyruvate carboxylase from corynebacterium glutamicum					
JOURNAL	Patent: US 6171833-A 1 09-JAN-2001;					
FEATURES	Location/Qualifiers					

source	1. .3621	/organism="unknown"	/mol_type="unassigned DNA"
ORIGIN			
Query Match	100.0%;	Score 3621;	DB 6; Length 3621;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 3621;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	TGGGGCGGGTTAGATCTGGGGGGTTATTTTCATTCATCTTGGCTGGAAGTCGTGAGG	60
Db	1	TGGGGCGGGTTAGATCTGGGGGGTTATTTTCATTCATCTTGGCTGGAAGTCGTGAGG	60
Qy	61	TCAGGGAGTGTTCGCCGAAACCATTCAGAGAGAAAACAAAACCGATGTTTGAATGGGGG	120
Db	61	TCAGGGAGTGTTCGCCGAAACCATTCAGAGAGAAAACAAAACCGATGTTTGAATGGGGG	120
Qy	121	AATCGGGGTTACATATAGGACGAGTGACTGCTATCACCCCTTGGGGTCTCTCTGTTG	180
Db	121	AATCGGGGTTACATATAGGACGAGTGACTGCTATCACCCCTTGGGGTCTCTCTGTTG	180
Qy	181	AAAGGAATAATTAATCTAGTGTGACTCACACATCTTCAACGCTTCAGCAATCAAAAG	240
Db	181	AAAGGAATAATTAATCTAGTGTGACTCACACATCTTCAACGCTTCAGCAATCAAAAG	240
Qy	241	ATCTTGTAGCAACCGCGCGGAAATCGCGTCCGTGCTTCCGTGCGAGCACTCGAAC	300
Db	241	ATCTTGTAGCAACCGCGCGGAAATCGCGTCCGTGCTTCCGTGCGAGCACTCGAAC	300
Qy	301	GGTGAGCCAGCGTAGCTATTTACCCCGGTGAAGATCGGGGATCATTCACCGCTCTTTT	360
Db	301	GGTGAGCCAGCGTAGCTATTTACCCCGGTGAAGATCGGGGATCATTCACCGCTCTTTT	360
Qy	361	GCTTCTGAAGCTGCCGATTTGTTACGAAAGCTCACAGTCAAGGCTTACTGGACATC	420
Db	361	GCTTCTGAAGCTGCCGATTTGTTACGAAAGCTCACAGTCAAGGCTTACTGGACATC	420
Qy	421	GATGAAATTTATCGGTGCAGCTAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC	480
Db	421	GATGAAATTTATCGGTGCAGCTAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC	480
Qy	481	TTCTGTCTGAAATATGCCAGCTTCCCGAGTGTGCGGAAACGAGCATTTACTTTTATT	540
Db	481	TTCTGTCTGAAATATGCCAGCTTCCCGAGTGTGCGGAAACGAGCATTTACTTTTATT	540
Qy	541	GGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAAACCCCGCG	600
Db	541	GGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAAACCCCGCG	600
Qy	601	AAGAAGCTGTCTGCGAGTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC	660
Db	601	AAGAAGCTGTCTGCGAGTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC	660
Qy	661	GTATAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAAGTTGCCGGTGGTGGC	720
Db	661	GTATAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAAGTTGCCGGTGGTGGC	720
Qy	721	GGACCGGTATCGGTTTGTGCTTCACTGTATGAGCTTCGCAATTAGCAACAGCA	780
Db	721	GGACCGGTATCGGTTTGTGCTTCACTGTATGAGCTTCGCAATTAGCAACAGCA	780
Qy	781	TCTCGTGAAGCTGAAGCGGCTTTCGCGCATGGCGGGTATATGTGCGAAGCTGCTGTGATT	840
Db	781	TCTCGTGAAGCTGAAGCGGCTTTCGCGCATGGCGGGTATATGTGCGAAGCTGCTGTGATT	840
Qy	841	AACCTTCAGCATTTGAAGTGAGATCTTTCGGGATCACACTGGAGAGTTGTACACCTT	900
Db	841	AACCTTCAGCATTTGAAGTGAGATCTTTCGGGATCACACTGGAGAGTTGTACACCTT	900
Qy	901	TATGAACGTGACTGTCTCACTGCAGGCTCGTCAACCAAAAAAGTTGTGCAAAATTCGCCAGCA	960
Db	901	TATGAACGTGACTGTCTCACTGCAGGCTCGTCAACCAAAAAAGTTGTGCAAAATTCGCCAGCA	960

2041 DB CCGCGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGCTCGCGCGCGTGTGT 2100  
2101 QY AAGGAAGCTGCGAGCTCCGCGGTGGACATCTTCGCGCATCTTCGACGGCTTTAAGCAGCTC 2160  
2101 DB AAGGAAGCTGCGAGCTCCGCGGTGGACATCTTCGCGCATCTTCGACGGCTTTAAGCAGCTC 2160  
2161 QY TCCAGATGCGTCCAGCAATCGACAGTCTCTGAGAGCAACACCGCGGTAGCGGAGGTG 2220  
2161 DB TCCAGATGCGTCCAGCAATCGACAGTCTCTGAGAGCAACACCGCGGTAGCGGAGGTG 2220  
2221 QY GCTATGGCTTATTCCTGGTGTATCTCTCTGATCCAAATGAAAGCTCTACACCTCGATTAC 2280  
2221 DB GCTATGGCTTATTCCTGGTGTATCTCTCTGATCCAAATGAAAGCTCTACACCTCGATTAC 2280  
2281 QY TACCTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGCTCACTCTTGGCCATTAAAGAT 2340  
2281 DB TACCTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGCTCACTCTTGGCCATTAAAGAT 2340  
2341 QY ATGGCTGTGTCTGCTCGCCAGCTGCGGTAAACCAAGCTGGTCAACGCACTGCGCGGTGAA 2400  
2341 DB ATGGCTGTGTCTGCTCGCCAGCTGCGGTAAACCAAGCTGGTCAACGCACTGCGCGGTGAA 2400  
2401 QY TTCGATCTGCCAGTGCACGTGCACACCGACACTGCGGGTGGCAGCTGGCACTTAC 2460  
2401 DB TTCGATCTGCCAGTGCACGTGCACACCGACACTGCGGGTGGCAGCTGGCACTTAC 2460  
2461 QY TTTGCTGCACTCAAGCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2520  
2461 DB TTTGCTGCACTCAAGCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2520  
2521 QY ACCACCTCCAGCCATCCCTGTCTGCCATTTGCTGCAATTTGCTGCAATTTGCTGCAATTTGCT 2580  
2521 DB ACCACCTCCAGCCATCCCTGTCTGCCATTTGCTGCAATTTGCTGCAATTTGCTGCAATTTGCT 2580  
2581 QY ACCGTTTGGAGCTCGAGCTGTTTCTGACCTGAGCGGTACTGGAAGCAGTGGCGGA 2640  
2581 DB ACCGTTTGGAGCTCGAGCTGTTTCTGACCTGAGCGGTACTGGAAGCAGTGGCGGA 2640  
2641 QY CTGTACTCTCCATTTGAGTCTGGAACCCCGAGCCCAACCGGTGCGGTCTACCGCCACGAA 2700  
2641 DB CTGTACTCTCCATTTGAGTCTGGAACCCCGAGCCCAACCGGTGCGGTCTACCGCCACGAA 2700  
2701 QY ATCCAGGCGGAGTGTTCGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
2701 DB ATCCAGGCGGAGTGTTCGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
2761 QY CGTTTTCGAACTCATCGAAGCAACTACGAGCGGTAAATGAGATGCTGGAGCGCCCAACC 2820  
2761 DB CGTTTTCGAACTCATCGAAGCAACTACGAGCGGTAAATGAGATGCTGGAGCGCCCAACC 2820  
2821 QY AAGGTCAACCCCATCTCCAAAGTGTGTCGCACTGCGACCTCCACTCGCTGCTGCTGCTGCTGCT 2880  
2821 DB AAGGTCAACCCCATCTCCAAAGTGTGTCGCACTGCGACCTCCACTCGCTGCTGCTGCTGCTGCT 2880  
2881 QY GTGGATCCAGCAGACTTTCGCGCATCCAAAGTACGACATCCAGACTCTGTCATC 2940  
2881 DB GTGGATCCAGCAGACTTTCGCGCATCCAAAGTACGACATCCAGACTCTGTCATC 2940  
2941 QY GGGTTCTCGCGGCGAGCTTGGTAACTCCAGTGGTGGCCAGAGCACTGCGCAC 3000  
2941 DB GGGTTCTCGCGGCGAGCTTGGTAACTCCAGTGGTGGCCAGAGCACTGCGCAC 3000  
3001 QY CCGGCACTGGAGGCGCTCCGAGGCGAGGCACTCTGACGGAAGTTCCTGAGGAAG 3060  
3001 DB CCGGCACTGGAGGCGCTCCGAGGCGAGGCACTCTGACGGAAGTTCCTGAGGAAG 3060  
3061 QY CAGGCGCACTCGACGCTGATGATTCGAAGGAACTGCGCAATAGCGCTCAACCGCTGCTG 3120  
3061 DB CAGGCGCACTCGACGCTGATGATTCGAAGGAACTGCGCAATAGCGCTCAACCGCTGCTG 3120  
3121 QY TTCCGAGGCCAACCGAAGTTCCTCGAGCACCCTGCGCGCTTGGGCAACCTCTGCG 3180

3121 DB TTCCGAGGCCAACCGAAGTTCCTCGAGCACCCTGCGCGCTTGGGCAACACCTCTGCG 3180  
3181 QY CTGGATGATCGTGAATTTCTTACCGCCTGCTGCGAGCGCGGAGACTTTGATTCGCGCTG 3240  
3181 DB CTGGATGATCGTGAATTTCTTACCGCCTGCTGCGAGCGCGGAGACTTTGATTCGCGCTG 3240  
3241 QY CCAGATGTCGCGCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300  
3241 DB CCAGATGTCGCGCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300  
3301 QY GGTATGCGCAATTTGTTGGCCAACTGCAACCGCCAGATCCGCCCAATGCGTGTGCTGAC 3360  
3301 DB GGTATGCGCAATTTGTTGGCCAACTGCAACCGCCAGATCCGCCCAATGCGTGTGCTGAC 3360  
3361 QY CGCTCGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
3361 DB CGCTCGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
3421 QY GTTGTCTGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3480  
3421 DB GTTGTCTGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3480  
3481 QY GCTGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 3540  
3481 DB GCTGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 3540  
3541 QY GTTGAACGCAAAATCGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600  
3541 DB GTTGAACGCAAAATCGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600  
3601 QY TTGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3621  
3601 DB TTGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3621

RESULT 2  
AR213173  
LOCUS AR213173 3621 bp DNA linear PAT 25-SEP-2002  
DEFINITION Sequence 1 from patent US 6403351.  
ACCESSION AR213173  
VERSION AR213173.1 GI:23310240  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 3621)  
AUTHORS Sinskey,A.J., Lessard,P.A. and Willis,L.B.  
TITLE Pyruvate carboxylase polypeptide from *Corynebacterium glutamicum*  
JOURNAL Patent: US 6403351-A 1 11-JUN-2002;  
FEATURES  
source Location/Qualifiers  
1..3621  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 3621; DB 6; Length 3621;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TGGGGCGGGGTAGATCTCTGGGGGGTTTATTTTCATTTCACTTTGGCTTGAAGTCGTGCGAGG 60  
Db 1 TGGGGCGGGGTAGATCTCTGGGGGGTTTATTTTCATTTCACTTTGGCTTGAAGTCGTGCGAGG 60  
Qy 61 TCAGGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAACAAACCGATGTTTGAATGGGGG 120  
Db 61 TCAGGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAACAAACCGATGTTTGAATGGGGG 120  
Qy 121 AATCGGGGGTTACGATCTAGTAGGACGAGTGTGCTATCACTTGGCGGTCTCTTGTG 180  
Db 121 AATCGGGGGTTACGATCTAGTAGGACGAGTGTGCTATCACTTGGCGGTCTCTTGTG 180  
Qy 181 AAGGAATATTAATCTAGTGTGCTGCTCACATCTTCAACGCTTCCAGCATTCAAAAAG 240

181 AAAGGAATTAATTAATCTAGTGTGACACACATCTTCAAGCTTCCAGCAATTCAAAAAG 240 Db  
241 ATCTTGGTAGCAAAACCGCGCGGAAATTCGCGGTCCGTCGAGCACTCGAAACC 300 Qy  
241 ATCTTGGTAGCAAAACCGCGCGGAAATTCGCGGTCCGTCGAGCACTCGAAACC 300 Db  
301 GGTGACGACCAAGTAGCTATTTACCCCGCTGAAGATCGGGGATCATTCACCGCTCTTTT 360 Qy  
301 GGTGACGACCAAGTAGCTATTTACCCCGCTGAAGATCGGGGATCATTCACCGCTCTTTT 360 Db  
361 GCTTCTGAAGCTCTCGGCAATGTGTACCGAAGGCTCACAGTCAAGCGTACTGTGACATC 420 Qy  
361 GCTTCTGAAGCTCTCGGCAATGTGTACCGAAGGCTCACAGTCAAGCGTACTGTGACATC 420 Db  
421 GATGAATTAATCGGTGAGTAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480 Qy  
421 GATGAATTAATCGGTGAGTAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480 Db  
481 TTCTGTCTGAAATGCCAGCTTGCCTGCGAGTGTGCGGAAACGGCAATTAATTTTAT 540 Qy  
481 TTCTGTCTGAAATGCCAGCTTGCCTGCGAGTGTGCGGAAACGGCAATTAATTTTAT 540 Db  
541 GGCCCAACCCAGAGTTCTTGATCTCACCGGTGATAAGTCTGCGCGGTAAACCGCGCG 600 Qy  
541 GGCCCAACCCAGAGTTCTTGATCTCACCGGTGATAAGTCTGCGCGGTAAACCGCGCG 600 Db  
601 AAGAAGCTGTCTGCGAGTTTGGCGGAATCCACCCCGAGCAAAATCGATGAGATC 660 Qy  
601 AAGAAGCTGTCTGCGAGTTTGGCGGAATCCACCCCGAGCAAAATCGATGAGATC 660 Db  
661 GTTAAAGCGCTGAAGCGCAGCTTACCCCATCTTGTGAAGCAGTTCGCGGTGTCG 720 Qy  
661 GTTAAAGCGCTGAAGCGCAGCTTACCCCATCTTGTGAAGCAGTTCGCGGTGTCG 720 Db  
721 GGAAGCGGTATGCGTTTGTGTTTCACTGTAGCTTGCAGAAATAGCAAGCA 780 Qy  
721 GGAAGCGGTATGCGTTTGTGTTTCACTGTAGCTTGCAGAAATAGCAAGCA 780 Db  
781 TCTGTGAAGCTGAAGCGCTTTCGGCGATGCGCGGTATATCGAAGTGTGAT 840 Qy  
781 TCTGTGAAGCTGAAGCGCTTTCGGCGATGCGCGGTATATCGAAGTGTGAT 840 Db  
841 AACCTCAGCATATGAAGTGCAGATCTTGGCGATCACACTGAGAAAGTTGACACCTT 900 Qy  
841 AACCTCAGCATATGAAGTGCAGATCTTGGCGATCACACTGAGAAAGTTGACACCTT 900 Db  
901 TATGAACGTGACTGCTCACTGACGCTGTCACCAAAAGTTGCGAAATTCGCGCAGCA 960 Qy  
901 TATGAACGTGACTGCTCACTGACGCTGTCACCAAAAGTTGCGAAATTCGCGCAGCA 960 Db  
961 CAGCAATTTGGATCCAGAACTGCGTGTATGCGGATGCGTAAAGTTCTGCGC 1020 Qy  
961 CAGCAATTTGGATCCAGAACTGCGTGTATGCGGATGCGTAAAGTTCTGCGC 1020 Db  
1021 TCCATTTGGTACCGAGGCGCGGAAACCGTGGAAATCTTGTGCGATGAAAGGCGAACCC 1080 Qy  
1021 TCCATTTGGTACCGAGGCGCGGAAACCGTGGAAATCTTGTGCGATGAAAGGCGAACCC 1080 Db  
1081 GTCTTTCATCGAAATGAACCCAGCTATCCAGTTCAGCACACCGTGAAGTCAAC 1140 Qy  
1081 GTCTTTCATCGAAATGAACCCAGCTATCCAGTTCAGCACACCGTGAAGTCAAC 1140 Db  
1141 GAGTGGACCTGGTGAAGCGCGAGATGCGTGTGCGTGTGCGAATTTGAAGGAATTG 1200 Qy  
1141 GAGTGGACCTGGTGAAGCGCGAGATGCGTGTGCGTGTGCGAATTTGAAGGAATTG 1200 Db  
1201 GGTCTGACCCCAAGTAAAGTCAAGACCCAGCTGAGCTGCGGATCCAGCAG 1260 Qy  
1201 GGTCTGACCCCAAGTAAAGTCAAGACCCAGCTGAGCTGCGGATCCAGCAG 1260 Db  
1261 GAAGATCCAAACAAAGCTTCCGCGCGAGATACCGGAATATCAACCGGTACCGCTACCA 1320 Qy  
1261 GAAGATCCAAACAAAGCTTCCGCGCGAGATACCGGAATATCAACCGGTACCGCTACCA 1320 Db

1321 GCGGAGCTGGCGTTCTGTCGCTGAGCTCAGTCCGTGGCGAAATCACCGCACAC 1380 Qy  
1321 GCGGAGCTGGCGTTCTGTCGCTGAGCTCAGTCCGTGGCGAAATCACCGCACAC 1380 Db  
1381 TTTGACTCCATGCTGTGAAATGACCTGCGGTGTTCCGACTTTGAAACTGCTGTGCT 1440 Qy  
1381 TTTGACTCCATGCTGTGAAATGACCTGCGGTGTTCCGACTTTGAAACTGCTGTGCT 1440 Db  
1441 CGTGCAAGCGCGCTGGCTGAGTTCAACGCTGCTGTGTTGCAACCAATTTGTTTC 1500 Qy  
1441 CGTGCAAGCGCGCTGGCTGAGTTCAACGCTGCTGTGTTGCAACCAATTTGTTTC 1500 Db  
1501 TTGCGTGGTGTGCTGCGGGAAGAGACTTCACTTCAAGCGCATGCGCAACCAATTCATT 1560 Qy  
1501 TTGCGTGGTGTGCTGCGGGAAGAGACTTCACTTCAAGCGCATGCGCAACCAATTCATT 1560 Db  
1561 GCGGATCACCGCACCTTTCAGGCTCCACCTGCTGATGATGAGCAGGAGCACTCCTG 1620 Qy  
1561 GCGGATCACCGCACCTTTCAGGCTCCACCTGCTGATGATGAGCAGGAGCACTCCTG 1620 Db  
1621 GATTACTTGGCAGATGTCACCGTGAACCAAGCTCATGCTGCGTCCAAAGGATGTTGCA 1680 Qy  
1621 GATTACTTGGCAGATGTCACCGTGAACCAAGCTCATGCTGCGTCCAAAGGATGTTGCA 1680 Db  
1681 GCTCTATCGATAGCTGCTTAACTCAAGGATCTGCCACTGCCAGCGGTTCCCGTGAC 1740 Qy  
1681 GCTCTATCGATAGCTGCTTAACTCAAGGATCTGCCACTGCCAGCGGTTCCCGTGAC 1740 Db  
1741 CGCTGAAGCAGCTTGGCCAGCGCGTTTCTGCTGATCTCCGTGAGCAGGAGCACTG 1800 Qy  
1741 CGCTGAAGCAGCTTGGCCAGCGCGTTTCTGCTGATCTCCGTGAGCAGGAGCACTG 1800 Db  
1801 CGAGTTACTGATACACTTCCGCGATGCAACAGCTTCTGCTGCGACCCGAGTCCGC 1860 Qy  
1801 CGAGTTACTGATACACTTCCGCGATGCAACAGCTTCTGCTGCGACCCGAGTCCGC 1860 Db  
1861 TCATTTCGCACTGAAGCTTGGCGAGGCGCTGCGCAAGCTGACTCTCTGAGCTTTGTC 1920 Qy  
1861 TCATTTCGCACTGAAGCTTGGCGAGGCGCTGCGCAAGCTGACTCTCTGAGCTTTGTC 1920 Db  
1921 GTGAGGCTTGGCGGCGGACCTACGATGCGGATGCTTCTCTCTTTGAGGATCCG 1980 Qy  
1921 GTGAGGCTTGGCGGCGGACCTACGATGCGGATGCTTCTCTCTTTGAGGATCCG 1980 Db  
1981 TGGGACAGCTTCGACAGCTGCGGAGCGGATGCGGATGTAACATTCAGATGCTGCTT 2040 Qy  
1981 TGGGACAGCTTCGACAGCTGCGGAGCGGATGCGGATGTAACATTCAGATGCTGCTT 2040 Db  
2041 GCGCGCGCAACACCGTGGGATACACCGGTACCCGCTCCGTCGCGCGGCTTTGTT 2100 Qy  
2041 GCGCGCGCAACACCGTGGGATACACCGGTACCCGCTCCGTCGCGCGGCTTTGTT 2100 Db  
2101 AAGGAGCTGCGAGCTCCGCGGTGACATCTTCGCGATCTTCGACGGGCTTAAAGAGTC 2160 Qy  
2101 AAGGAGCTGCGAGCTCCGCGGTGACATCTTCGCGATCTTCGACGGGCTTAAAGAGTC 2160 Db  
2161 TCCAGATGCGTTCAGCAATCGAGCTCTGAGACCAACACCGGATGAGCGAGTG 2220 Qy  
2161 TCCAGATGCGTTCAGCAATCGAGCTCTGAGACCAACACCGGATGAGCGAGTG 2220 Db  
2221 GCTATGGCTTATCTGGTGATCTCTCATCAAAATGAAAGCTCTACACCTGAGATTAC 2280 Qy  
2221 GCTATGGCTTATCTGGTGATCTCTCATCAAAATGAAAGCTCTACACCTGAGATTAC 2280 Db  
2281 TACCTTAAAGATGGCAGAGGATCGTCAAGTCTGCGCTCACTCTTGGCATTAAAGAT 2340 Qy  
2281 TACCTTAAAGATGGCAGAGGATCGTCAAGTCTGCGCTCACTCTTGGCATTAAAGAT 2340 Db  
2341 ATGGCTGTGCTTTCGCGCGAGCTGCGTAAACAGCTGGTCAACCGACTGCGCGTCAA 2400 Qy  
2341 ATGGCTGTGCTTTCGCGCGAGCTGCGTAAACAGCTGGTCAACCGACTGCGCGTCAA 2400 Db



Qy	2401	TTCCGATCTGCCAGTGCACGTGCACACCCACGACACTCTCGCGGTGGCCAGCTGGCAACCTTAC	2460
Db	2401	TTCCGATCTGCCAGTGCACGTGCACACCCACGACACTCTCGCGGTGGCCAGCTGGCAACCTTAC	2460
Qy	2461	TTTGTCTGCAGCTCAAGCTGTGTGCAGATGCTGTGAACGGTGTCTTCGCGACCACTGTGTCTGGC	2520
Db	2461	TTTGTCTGCAGCTCAAGCTGTGTGCAGATGCTGTGAACGGTGTCTTCGCGACCACTGTGTCTGGC	2520
Qy	2521	ACCACTCTCCAGCCATCCCTGTCTGCCATTTGTCTGTCTCATTTCCGCGCACACCCCGTTCGCGAT	2580
Db	2521	ACCACTCTCCAGCCATCCCTGTCTGCCATTTGTCTGTCTCATTTCCGCGCACACCCCGTTCGCGAT	2580
Qy	2581	ACCGGTTTGAGCCTCGAGGCTGTCTTGACCTCGAGCCGTACTTGGGAGCAGTGGCGGGA	2640
Db	2581	ACCGGTTTGAGCCTCGAGGCTGTCTTGACCTCGAGCCGTACTTGGGAGCAGTGGCGGGA	2640
Qy	2641	CTGTACTCTGCCATTTGAGTCTGGAAACCCCAAGGCCCAACCGGTTCGCGTCTACCGCCACGAA	2700
Db	2641	CTGTACTCTGCCATTTGAGTCTGGAAACCCCAAGGCCCAACCGGTTCGCGTCTACCGCCACGAA	2700
Qy	2701	ATCCGAGGCGGACAGTTGTCTCAACCTCGCTGTCACAGGCCACCGGACCTTGGGCTTTCGGGAT	2760
Db	2701	ATCCGAGGCGGACAGTTGTCTCAACCTCGCTGTCACAGGCCACCGGACCTTGGGCTTTCGGGAT	2760
Qy	2761	CGTTTTCGAACTCATTCGAAGCAACAACCTAGCAGCCGGTTAATCAGATGCTTGGGACGCCCAACC	2820
Db	2761	CGTTTTCGAACTCATTCGAAGCAACAACCTAGCAGCCGGTTAATCAGATGCTTGGGACGCCCAACC	2820
Qy	2821	AAGGTCAACCCCATCTCTCAAGGTTGTTGGGCACTCTCGCACTCCCACTCTGTTGTGTGGGCT	2880
Db	2821	AAGGTCAACCCCATCTCTCAAGGTTGTTGGGCACTCTCGCACTCCCACTCTGTTGTGTGGGCT	2880
Qy	2881	GTGGATCCAGCAGACTTTGTCTGCCGATCCCAAAAGTACGACACTCCCAGACTCTGTCTATC	2940
Db	2881	GTGGATCCAGCAGACTTTGTCTGCCGATCCCAAAAGTACGACACTCCCAGACTCTGTCTATC	2940
Qy	2941	GCCTTCCTCGCGCGAGCTTGGTAAACCTCCAGGTGGCTTGGCCAGAGCCCACTCTCGCAC	3000
Db	2941	GCCTTCCTCGCGCGAGCTTGGTAAACCTCCAGGTGGCTTGGCCAGAGCCCACTCTCGCAC	3000
Qy	3001	CGCGCATGGAAGCCGCTCCGAGGCAAGGCACTCTGACGGAGTTCCTGAGGAGAG	3060
Db	3001	CGCGCATGGAAGCCGCTCCGAGGCAAGGCACTCTGACGGAGTTCCTGAGGAGAG	3060
Qy	3061	CAGCGCACCTCGACGCTGTATGATTTCCAAGGAACGTGCGCAATAGCTCAACCGCTCTGTG	3120
Db	3061	CAGCGCACCTCGACGCTGTATGATTTCCAAGGAACGTGCGCAATAGCTCAACCGCTCTGTG	3120
Qy	3121	TTCCGGAAGCCACCGAAGAGTTCTCTGAGCACCGTTCGCGCTTCGGGCAACACTCTGTGGC	3180
Db	3121	TTCCGGAAGCCACCGAAGAGTTCTCTGAGCACCGTTCGCGCTTCGGGCAACACTCTGTGGC	3180
Qy	3181	CTGGATGATCGTGAATTTCTTACGGGCTGTGTCCAAGCCGCGAGACTTTTGATCCGGCTGTG	3240
Db	3181	CTGGATGATCGTGAATTTCTTACGGGCTGTGTCCAAGCCGCGAGACTTTTGATCCGGCTGTG	3240
Qy	3241	CCAGATGTGGCACCCCACTGCTTGTTCGCTTGAATTCGATCTCTGAGCCAGACGATAAG	3300
Db	3241	CCAGATGTGGCACCCCACTGCTTGTTCGCTTGAATTCGATCTCTGAGCCAGACGATAAG	3300
Qy	3301	GGTATGCGCAATGTTGTGGCCAAACGTCAAAGGCCAGATCCGCCCAATTCGCTGTGCTGTC	3360
Db	3301	GGTATGCGCAATGTTGTGGCCAAACGTCAAAGGCCAGATCCGCCCAATTCGCTGTGCTGTC	3360
Qy	3361	CGCTCCGTTGAGTCTGCACCGCAACCGCAAAAAGGCAGATTCTCTCAAACAGGGCCAT	3420
Db	3361	CGCTCCGTTGAGTCTGCACCGCAACCGCAAAAAGGCAGATTCTCTCAAACAGGGCCAT	3420
Qy	3421	GTTCCTGCACCAATTCGCTGGTGTGTCAACCGTGACTGCTGCTGAAGTGTATGAGGTCAAG	3480
Db	3421	GTTCCTGCACCAATTCGCTGGTGTGTCAACCGTGACTGCTGCTGAAGTGTATGAGGTCAAG	3480
Qy	3481	GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCAGCTGCTTCT	3540

Db	3481	GCTGGAGATGCACTGCGCAATCATCGAGGCTATGAAGATGGAACCAACAATCACTGCTTCT	3540
Qy	3541	GTTGACGGCAAAATCGATCGCGTGTGTGGTTCCTGCTGCAACGAAGTGGAAAGGTGGCGAC	3600
Db	3541	GTTGACGGCAAAATCGATCGCGTGTGTGGTTCCTGCTGCAACGAAGTGGAAAGGTGGCGAC	3600
Qy	3601	TTGATCGTCGTCGTTTCCTAA	3621
Db	3601	TTGATCGTCGTCGTTTCCTAA	3621
RESULT 3			
LOCUS	AF038548		
DEFINITION	Corynebacterium glutamicum pyruvate carboxylase (pyc) gene, complete cds.		
ACCESSION	AF038548		
VERSION	AF038548.1	GI:2708716	
KEYWORDS	Corynebacterium glutamicum		
SOURCE	Corynebacterium glutamicum		
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.		
REFERENCE	1 (bases 1 to 3637)		
AUTHORS	Koffas, M.A., Ramamoorthi, R., Pine, W.A., Sinskey, A.J. and Stephanopoulos, G.		
TITLE	Sequence of the Corynebacterium glutamicum pyruvate carboxylase gene		
JOURNAL	Appl. Microbiol. Biotechnol. 50 (3), 346-352 (1998)		
MEDLINE	Appl. Microbiol. Biotechnol. 50 (3), 346-352 (1998)		
PUBMED	93019028		
REFERENCE	2 (bases 1 to 3637)		
AUTHORS	Koffas, M.A.G., Ramamoorthi, R., Pine, W.A., Sinskey, A.J. and Stephanopoulos, G.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-DEC-1997) Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA		
FEATURES	Location/Qualifiers		
source	1..3637		
	/organism="Corynebacterium glutamicum"		
	/mol_type="genomic DNA"		
	/strain="21253"		
	/db_xref="taxon:1718"		
gene	199..3621		
	/gene="pyc"		
CDS	199..3621		
	/gene="pyc"		
	/codon_start=1		
	/transl_table=11		
	/product="pyruvate carboxylase"		
	/protein_id="AAB92588.1"		
	/db_xref="GI:2708717"		
	/translations="MSTHTSSSTTAPFKTILVANRGVIAPRAALETGTAATVAIYPR		
	EDRGSFERSFASAVRIGTSGSPVKAVLDIDELIGAKKVKADAIIPGYGFLSNQGL		
	ARCEANRITFPTGDEVLDTGDKRAVTAARKAGLPVLAESTPSSNIDETIKSAGSL		
	QYPIFYKAVAGGGRGMPFASPDRLKLAETAREEAAGFGDGVYVERAVINPOH		
	IYQIILGDTHGEVHLYERDCSLQRHQKVVEIAPAQLDPELRDRI CADAVKFCRSI		
	GYOGAKTEFLPEDEKGNHVIEMNPRIQVEHTEETVEDLVKAQMLRAAGATKEL		
	GLTQDKTIGLGAALQCRITTEDPNGFRPDTGVTITAVRSPGGAVRLDGAALQGGIT		
	AHFDLSLMTCKGSPDEFAVARQALAEFTVSGVATNIGFLRALLREDFDKRLTA		
	TGTIADPHLLQAPADDEGRLLDALVDTVKNKPGVRPKDVAAPIDKLPNIKDLPL		
	PGRDRRLKQAPAAFDREQLDALVDTTFDRAHQSLATVRSFALKPAAEVAVA		
	KLTPLELSVEAGAGSYDVAMRFLFEDPDWRLDELREAMPNVIQMLRGNTVGGDL		
	KYPDSVCRFAKKEASGVDTIFRFDALNDVSRMPAIDAVLENTVAEVAAMAYSGD		
	SDNEKLYITDLYLKMAREITVKSAGILAIKDMAGLIRPAATVKLVTALRRERDPLVH		
	VHSTDAGGGLATVFAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRDGTLS		
	LEAVSDLPYWEAVRGVILPFESGTPGCTGVYRHEIPGQSLNLRQAQATGLADRF		
	ELIEDNYAANVEMGLPRTKTPSKVVGDLALHLVGVGVDPADPADPKQYDIPDSVI		
	AFLRGEUNGPWGPMPKPTVALEGRSEKAPLITEVEPEEQALDDADDKERNLNLR		
	LDFLGEKTEEFLEFRNGFNSTALDDREFFLPGYVEGRTLIRLDVTRFLPLLDALISE		
	PDDKGMENYVANGRIQPNRVRDSVESVTALEKADSSNKGHVAAPFAGVTVTTVA		

EGDEVKAGDAVAIIIEAMKMEATITASVDKIDRVVVVPAATKVEGGDLIVVVS"

ORIGIN			
Query Match			
Best Local Similarity 100.0%; Score 3621; DB 1; Length 3637;			
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	TGGGGGGGGTATAGATCTCTGGGGGGTATTTTTCATTCACTTTGGCTTGAAGTCGTGCAAG 60	
Db	1	TGGGGGGGGTATAGATCTCTGGGGGGTATTTTTCATTCACTTTGGCTTGAAGTCGTGCAAG 60	
Qy	61	TCAGGGAGTGTTCGGGAAACATTGAGAGGAAACAAACCCGATGTTGATTTGGGG 120	
Db	61	TCAGGGAGTGTTCGGGAAACATTGAGAGGAAACAAACCCGATGTTGATTTGGGG 120	
Qy	121	AATCGGGGTTACGATACAGTACGAGTACGCTATCACCCCTTGGCGGTCTCTTTGTTG 180	
Db	121	AATCGGGGTTACGATACAGTACGAGTACGCTATCACCCCTTGGCGGTCTCTTTGTTG 180	
Qy	181	AAAGGAATAATTAATCTAGTGTGATCACAATCTTCAACCGTTCAGCATTCAAAAG 240	
Db	181	AAAGGAATAATTAATCTAGTGTGATCACAATCTTCAACCGTTCAGCATTCAAAAG 240	
Qy	241	ATCTTGTGTAGCAACCGCGCGAAATCGCGTCCGTCTTTCGCTGAGCATTCGAAACC 300	
Db	241	ATCTTGTGTAGCAACCGCGCGAAATCGCGTCCGTCTTTCGCTGAGCATTCGAAACC 300	
Qy	301	GGTGAGCCAGGTAGTATTTACCCCGTGAAGATCGGGGATCATTCACCGCTCTTTT 360	
Db	301	GGTGAGCCAGGTAGTATTTACCCCGTGAAGATCGGGGATCATTCACCGCTCTTTT 360	
Qy	361	GCTTCTCAAGTGTCCGATTTGTACCGAAGGTCAACAGTCAAGCGTACCTTGAATC 420	
Db	361	GCTTCTCAAGTGTCCGATTTGTACCGAAGGTCAACAGTCAAGCGTACCTTGAATC 420	
Qy	421	GATGAATATTCGTGTGAGTAAAGATTAAGAGATGCCATTTACCGGGATACGGC 480	
Db	421	GATGAATATTCGTGTGAGTAAAGATTAAGAGATGCCATTTACCGGGATACGGC 480	
Qy	481	TTCTGTCTGAATATGCCAGTTCGCCGAGTGTGCGGAAACAGGATTAATTTAT 540	
Db	481	TTCTGTCTGAATATGCCAGTTCGCCGAGTGTGCGGAAACAGGATTAATTTAT 540	
Qy	541	GGCCAAACCCAGAGTTCGTATCTCACCGGTGATAGTCTCGCGGGTAAACCGCGCG 600	
Db	541	GGCCAAACCCAGAGTTCGTATCTCACCGGTGATAGTCTCGCGGGTAAACCGCGCG 600	
Qy	601	AAGAGGTGTCTGCCAGTTCGCGGAATCCACCCGAGCAAAACATCGATGATC 660	
Db	601	AAGAGGTGTCTGCCAGTTCGCGGAATCCACCCGAGCAAAACATCGATGATC 660	
Qy	661	GTTAAAGCGTGAAGGCCAGACTTACCCATCTTTGTGAAGGAGTTCGCGGTGGTGGC 720	
Db	661	GTTAAAGCGTGAAGGCCAGACTTACCCATCTTTGTGAAGGAGTTCGCGGTGGTGGC 720	
Qy	721	GAGCGGTATGCGTTTGTCTACCTGATGAGTTCGCAATTTAGCAACAGAGCA 780	
Db	721	GAGCGGTATGCGTTTGTCTACCTGATGAGTTCGCAATTTAGCAACAGAGCA 780	
Qy	781	TCTGTGAAGTGAAGCGGCTTTGGCGATGGCGGTATATGTCGAACGTGTGAT 840	
Db	781	TCTGTGAAGTGAAGCGGCTTTGGCGATGGCGGTATATGTCGAACGTGTGAT 840	
Qy	841	AACCTTACGATATTAAGTGCAGATCCTTGGGATCACCTGAGAGTGTACACCT 900	
Db	841	AACCTTACGATATTAAGTGCAGATCCTTGGGATCACCTGAGAGTGTACACCT 900	
Qy	901	TATGAACGTGATGCTCACTGAGCGTCTGATCCAAAGTGTGCAAAATTCGCCAGCA 960	
Db	901	TATGAACGTGATGCTCACTGAGCGTCTGATCCAAAGTGTGCAAAATTCGCCAGCA 960	
Qy	961	CAGCATTTGGATCCGAACCTGGTATCGCATTTGTGGGATGAGTAAAGTTCGCGC 1020	

Db	961	CAGCATTTGGATCCGAACCTGGTATCGCATTTGTGGGATGAGTAAAGTTCGCGC 1020	
Qy	1021	TCATTGGTTTACCAGGGCGCGGAAACCGTGAATTTCTTGGTGCATGAAAGGGCAACAC 1080	
Db	1021	TCATTGGTTTACCAGGGCGCGGAAACCGTGAATTTCTTGGTGCATGAAAGGGCAACAC 1080	
Qy	1081	GTCTTCATCGAATGAACCAAGTATCCAGTTCAGCAGCAGTGCAGTGAAGAGTCACC 1140	
Db	1081	GTCTTCATCGAATGAACCAAGTATCCAGTTCAGCAGCAGTGCAGTGAAGAGTCACC 1140	
Qy	1141	GAGTGTGACCTGTGTGAAGCGCGAGATGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCT 1200	
Db	1141	GAGTGTGACCTGTGTGAAGCGCGAGATGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCT 1200	
Qy	1201	GGTGTGACCCAGATTAAGATCAAGACCCAGGTCAGCTGCGTGGCGGATCAACACG 1260	
Db	1201	GGTGTGACCCAGATTAAGATCAAGACCCAGGTCAGCTGCGTGGCGGATCAACACG 1260	
Qy	1261	GAAGATCAAAACAAAGGCTTCCGCCAGATACCGGAATATATCACCGGTACCGCTCACCA 1320	
Db	1261	GAAGATCAAAACAAAGGCTTCCGCCAGATACCGGAATATATCACCGGTACCGCTCACCA 1320	
Qy	1321	GGCGGAGTGGCGTTCGTCTGAGCGTGCAGCTCAGCTGCGTGGCGGATCAACGCGCAC 1380	
Db	1321	GGCGGAGTGGCGTTCGTCTGAGCGTGCAGCTCAGCTGCGTGGCGGATCAACGCGCAC 1380	
Qy	1381	TTTGACTCCATGCTGTGATAAATGACCTGCGGTGGTTCGACTTTGAAACTGCTGTGCT 1440	
Db	1381	TTTGACTCCATGCTGTGATAAATGACCTGCGGTGGTTCGACTTTGAAACTGCTGTGCT 1440	
Qy	1441	CGTGACAGCGCGGTGGCTGAGTTCACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1500	
Db	1441	CGTGACAGCGCGGTGGCTGAGTTCACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1500	
Qy	1501	TTGGTGTGCTGCTGCGGAGAGACTTCACTTCCAAAGCGCATCGCACCGGATTCATT 1560	
Db	1501	TTGGTGTGCTGCTGCGGAGAGAGACTTCACTTCCAAAGCGCATCGCACCGGATTCATT 1560	
Qy	1561	CGCGATACCCGCGCTCTTTCAGGTCACCTGCTGATGATGAGCAGGAGCGCATCTG 1620	
Db	1561	CGCGATACCCGCGCTCTTTCAGGTCACCTGCTGATGATGAGCAGGAGCGCATCTG 1620	
Qy	1621	GATTACTTGGCAGATGTCAACCGTGAACAGGCTCATGCTGCTGCTGCTGCTGCTGCTGCT 1680	
Db	1621	GATTACTTGGCAGATGTCAACCGTGAACAGGCTCATGCTGCTGCTGCTGCTGCTGCTGCT 1680	
Qy	1681	GCTCTATCGATTAAGTGCCTAAACATCAAGGATCTGCCACTGCCACCGGTTCCCGTGAC 1740	
Db	1681	GCTCTATCGATTAAGTGCCTAAACATCAAGGATCTGCCACTGCCACCGGTTCCCGTGAC 1740	
Qy	1741	CGCTTGAAGCAGTTCGCCCGCGGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800	
Db	1741	CGCTTGAAGCAGTTCGCCCGCGGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800	
Qy	1801	GCAGTTACTGTATACCACTTCCGCGATGCAACACAGTCTTTCGCTGCTGCTGCTGCTGCTGCT 1860	
Db	1801	GCAGTTACTGTATACCACTTCCGCGATGCAACACAGTCTTTCGCTGCTGCTGCTGCTGCTGCT 1860	
Qy	1861	TCATTTCGATGAGCTTGGCGGAGAGCGCTGCGCAAGCTGACTCTCTGAGCTTTTGTCC 1920	
Db	1861	TCATTTCGATGAGCTTGGCGGAGAGCGCTGCGCAAGCTGACTCTCTGAGCTTTTGTCC 1920	
Qy	1921	GTGGAGGCTTGGCGGCGGCGGACCTTACGATGTGGCGATGCGTTCCTCTTTGAGGATCG 1980	
Db	1921	GTGGAGGCTTGGCGGCGGCGGACCTTACGATGTGGCGATGCGTTCCTCTTTGAGGATCG 1980	
Qy	1981	TGGGACAGCTCGAGAGTTCGCCGAGGCGGATGCGCAATGTAACATTCAGATGCTGCTT 2040	
Db	1981	TGGGACAGCTCGAGAGTTCGCCGAGGCGGATGCGCAATGTAACATTCAGATGCTGCTT 2040	
Qy	2041	CGCGGCGGCAACACCGTGGGATACACCGGATACACCGGATACACCGGATACACCGGATAC 2100	
Db	2041	CGCGGCGGCAACACCGTGGGATACACCGGATACACCGGATACACCGGATACACCGGATAC 2100	

```
QY 2101 AAGGAGCTGCGAGCTCGGCGTGAGCATCTTCGCGCATCTTCGACGCGCTTAAAGAGCTC 2160
Db 2101 AAGGAGCTGCGAGCTCGGCGTGAGCATCTTCGCGCATCTTCGACGCGCTTAAAGAGCTC 2160
QY 2161 TCCAGATGCGTCCAGCAATCGAGCGAGTCTCGAGACCAACACCGCGGTAGCGAGGTG 2220
Db 2161 TCCAGATGCGTCCAGCAATCGAGCGAGTCTCGAGACCAACACCGCGGTAGCGAGGTG 2220
QY 2221 GCTATGGCTTATCTCGGTGATCTCTCTGATCCAAATGAAAGCTCTACACCTGGATTAC 2280
Db 2221 GCTATGGCTTATCTCGGTGATCTCTCTGATCCAAATGAAAGCTCTACACCTGGATTAC 2280
QY 2281 TACCTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGCTCACAATCTTGGCCATTAAAGAT 2340
Db 2281 TACCTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGCTCACAATCTTGGCCATTAAAGAT 2340
QY 2341 ATGGCTGTCTGCTTCGCCAGCTGCGGTAAACAGCTGCTCAGCACTGCGCACTGCGCGTAA 2400
Db 2341 ATGGCTGTCTGCTTCGCCAGCTGCGGTAAACAGCTGCTCAGCACTGCGCACTGCGCGTAA 2400
QY 2401 TTGATCTGCCAGTGCAGTGCACACCCACGACACTGCGGGTGCCAGCTGGCAACCTAC 2460
Db 2401 TTGATCTGCCAGTGCAGTGCACACCCACGACACTGCGGGTGCCAGCTGGCAACCTAC 2460
QY 2461 TTGATCTGCCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 2520
Db 2461 TTGATCTGCCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 2520
QY 2521 ACCACTCTCCAGCCATCCTGTCTGCCATTTGCTGCAATTCGGGACACACCCGTGCGGAT 2580
Db 2521 ACCACTCTCCAGCCATCCTGTCTGCCATTTGCTGCAATTCGGGACACACCCGTGCGGAT 2580
QY 2581 ACCGTTTGGAGCTCGAGCTGTTTCTGACCTCGAGCGCTACTGGAGAGAGTGGCGGA 2640
Db 2581 ACCGTTTGGAGCTCGAGCTGTTTCTGACCTCGAGCGCTACTGGAGAGAGTGGCGGA 2640
QY 2641 CTGTACTGCTGCAATTTGAGTCTGGAACCCAGGCCCAACCGGTGCGGTCTACCGCCAGAA 2700
Db 2641 CTGTACTGCTGCAATTTGAGTCTGGAACCCAGGCCCAACCGGTGCGGTCTACCGCCAGAA 2700
QY 2701 ATCCAGGCGGAGTGTGTCAACTGCTGGTGACAGGCCACCGCACTGGGCTTGGGAT 2760
Db 2701 ATCCAGGCGGAGTGTGTCAACTGCTGGTGACAGGCCACCGCACTGGGCTTGGGAT 2760
QY 2761 CGTTTCGAATCATCGAAGACAATACGAGCGCTTAATGAGATGCTGGGACGCGCCAAAC 2820
Db 2761 CGTTTCGAATCATCGAAGACAATACGAGCGCTTAATGAGATGCTGGGACGCGCCAAAC 2820
QY 2821 AAGGTCAACCCATCTCCAGGTGTTGGCGACCTCGCACTCCACCTCGTTGGTGGCGGT 2880
Db 2821 AAGGTCAACCCATCTCCAGGTGTTGGCGACCTCGCACTCCACCTCGTTGGTGGCGGT 2880
QY 2881 GTGGATCCAGCAGACTTGTCTCGCATCCAAAGTACGATCCAGATCTGTCTATC 2940
Db 2881 GTGGATCCAGCAGACTTGTCTCGCATCCAAAGTACGATCCAGATCTGTCTATC 2940
QY 2941 GGGTTCTCTCGCGGCGAGCTTGGTAAACCTCCAGGTGGTGGCCAGAGCCACTGCGCAC 3000
Db 2941 GGGTTCTCTCGCGGCGAGCTTGGTAAACCTCCAGGTGGTGGCCAGAGCCACTGCGCAC 3000
QY 3001 CGGCACTGGAAGGCGCTCGAAGCAGCACTCTGACGAGATGCTTGGAGAGAG 3060
Db 3001 CGGCACTGGAAGGCGCTCGAAGCAGCACTCTGACGAGATGCTTGGAGAGAG 3060
QY 3061 CAGGCGCACTCGACGCTGATGTTCCAAAGAACCTGCGCAATAGCCTCAACCGCTGCTG 3120
Db 3061 CAGGCGCACTCGACGCTGATGTTCCAAAGAACCTGCGCAATAGCCTCAACCGCTGCTG 3120
QY 3121 TTCCCGAAGCCCAACCGAAGAGTCTCTCGAGCAGCGTTCGGGCAACACCTCTGG 3180
Db 3121 TTCCCGAAGCCCAACCGAAGAGTCTCTCGAGCAGCGTTCGGGCAACACCTCTGG 3180
```

```
QY 3181 CTGGATGATCGTGAATTTCTTACGCGCTGGTTCGAAGCCGCGAGACTTTGATCGCGTG 3240
Db 3181 CTGGATGATCGTGAATTTCTTACGCGCTGGTTCGAAGCCGCGAGACTTTGATCGCGTG 3240
QY 3241 CCAGATGTCGCGACCCCACTGCTTGGCTGGATGGATCTCTGAGCCAGACGATAAG 3300
Db 3241 CCAGATGTCGCGACCCCACTGCTTGGCTGGATGGATCTCTGAGCCAGACGATAAG 3300
QY 3301 GGTATGCGCAATTTGTGTGCCAACTCAACGCGCCAGATCCGCCCAATCGTGTGGTGAAC 3360
Db 3301 GGTATGCGCAATTTGTGTGCCAACTCAACGCGCCAGATCCGCCCAATCGTGTGGTGAAC 3360
QY 3361 CGCTCGCTTGAATCTGTCACGCAACCCGAGAAAGGAGATTCCTCAACAAAGGCCAT 3420
Db 3361 CGCTCGCTTGAATCTGTCACGCAACCCGAGAAAGGAGATTCCTCAACAAAGGCCAT 3420
QY 3421 GTTGTGTCACCACTCGCTGCTTGTGTCACCGTGAATCTTGTGTAAGGTGATGAGTCAAG 3480
Db 3421 GTTGTGTCACCACTCGCTGCTTGTGTCACCGTGAATCTTGTGTAAGGTGATGAGTCAAG 3480
QY 3481 GTGGAGATGCAATGCGCAATCATCGAGCTATGAAGATGGAAGCAACATCACTGCTTCT 3540
Db 3481 GTGGAGATGCAATGCGCAATCATCGAGCTATGAAGATGGAAGCAACATCACTGCTTCT 3540
QY 3541 GTTGAAGGCAAAATCGATCGCTTGTGTTCTGCTGCAACGAAGGTGGAAGTGGCGAC 3600
Db 3541 GTTGAAGGCAAAATCGATCGCTTGTGTTCTGCTGCAACGAAGGTGGAAGTGGCGAC 3600
QY 3601 TTGATCGTCTGCTGCTTCTTAA 3621
Db 3601 TTGATCGTCTGCTGCTTCTTAA 3621
```

## RESULT 4

AP005276 332050 bp DNA linear BCT 08-AUG-2002  
LOCUS Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 3/10.

ACCESSION AP005276 BAO000036

VERSION AP005276.1 GI:21323419

## KEYWORDS

SOURCE Corynebacterium glutamicum ATCC 13032

## ORGANISM

Corynebacterium glutamicum ATCC 13032  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriaceae; Corynebacterium.

## REFERENCE

1. Nakagawa, S.  
Complete genomic sequence of Corynebacterium glutamicum ATCC 13032

Unpublished

2. (bases 1 to 332050)

## AUTHORS

Direct Submission

## JOURNAL

Nakagawa, S.

## TITLE

Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.

Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,

Tokyo 194-8533, Japan (E-mail: snakagawa@xanagen.com,

Tel: 81-44-829-3031, Fax: 81-44-813-1651)

This sequence is conducted by collaboration of Kyowa Hakko Kogyo

Co. Ltd. And Kitasato University.

## FEATURES

## source

1. .332050

/organism="Corynebacterium glutamicum ATCC 13032"

/mol\_type="genomic DNA"

/strain="ATCC 13032"

/db\_xref="taxon:196627"

/notes="ATCC 13032"

314. .2518

/gene="Cgl10654"

314. .2518

/gene="Cgl10654"

/note="PF02554:Carbon starvation protein Csta"

/codon\_start=1

/transl\_table=11

/product="Carbon starvation protein, predicted membrane

```

protein_id="BAB98047.1"
/db_xref="GI:21323420"
/translation="MKSPKMGFTARVGLLVFGVIAAVNGMAIAFSRGETINSVWLVL
AAVGSYIIAFSYRIEYKVKPKQQRATPAAYNDGKDYPTDRVLFGHFAAIA
GAGPLVPMQAQMGVLPGLTIIILGIVNAGAVQDYLVLVSTRRRGRSLQGVNDEM
GTGGAGATLATSIIIIIAVLIVNAGADSPGVPSTIMTPIALFNGVLYRL
RGRVTEVSLIPGVALIIVAGVWADTSWGVNFTWSTKTLALALIGYMAAILP
VHLLIAPRYLSFPMKGVIGILLAVGILPARPEVOMPSVTSFALGNGPVSGLSFPF
LFTIACALSGFHALISGTPPKLVKESQRMGLYCGGMLNESFVAMALITAVILD
RHLYFMNAPLALGTGGPATAEWNISIGLTGADITPEQLSEASVSGESTVISRTGG
APTLAFMGSEIILSGFTGGGAKAFWFHFAIMEALFITLTVDAGTRVAFWSTIYK
VGLRFPDPSWTVGNWISFVFCALWGAIIILMGVTDPLGGINVLFPFLGALQALAA
IALALRVVVKGLVKWAWIPAVPLAWDLIVMTASWOKIHFSPDPAIGYWAQNAFR
DAKSQGLTFEGAASPEADAVIRNTMIQGISILFVILVVGGAIAVCVICSIRAR
AAGTPLETHEEPTSEFPAPGFLASSRDKVEQAWDBERYFGGAPVSSGGH"
2522..2725
/gene="Cg10655"
CDS
2522..2725
/gene="Cg10655"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB98048.1"
/db_xref="GI:21323421"
/translation="MMALTHALWKIPRAVWVYLTLMGMTAYSKYVHVLKHHHPAPI
PTREYWRAXYADQDANPGARCC"
complement(3092..3703)
/gene="Cg10656"
CDS
complement(3092..3703)
/gene="Cg10656"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB98049.1"
/db_xref="GI:21323422"
/translation="MSNVINLVMQASHOVNLIPLSSDATDQILATFFSLCCSQ
NRNVLVSWGVFSEERVMEIKLTHRSVSKSPFEGITSGWOTENRCTPIGNWPDLC
HTRGRANWSQHSGSLRCHRCIVDAVDHDLVSRLNLMWGNASDLVSQVLLVRELFR
RTGVAHFVNVHNVNTPSQRTNVDAHDHNLKSNVGHCKSCVIL"
3750..4706
/gene="Cg10657"
CDS
3750..4706
/gene="Cg10657"
/codon_start=1
/transl_table=11
/product="Uncharacterized protein involved in propionate
catabolism"
/protein_id="BAB98050.1"
/db_xref="GI:21323423"
/translation="MNTIYQAIQALHTTTATRSQRKGAISWKAFAPAFAGKMSIE
AVDRAMEGAPSPWEGEDGVIALLSGLDHYITPLPABGEAKRAIDTYTKGSHA
EYQSQPIDLARSNGEKLAAGLDLRVDOSIVLHTSHHTYVIGTSGNDPOKFPDPA
REILDHIMTIFVALEDRAWHRSYAPERAHRETIELWNKISIVDEPENTRYHS
VDPAAFGARAVITFKDGVTVDELAADAHPLGARPFARSQYIQRFLAEGVVSSE
KEQDRFLDAQORTHELELDELSELNIELDADILAKAPVPEGLF"
4715..5629
/gene="Cg10658"
CDS
4715..5629
/gene="Cg10658"
/codon_start=1
/transl_table=11
/product="PEP phosphonmutase and related enzymes"
/protein_id="BAB98051.1"
/db_xref="GI:21323424"
/translation="MFSSAVAPTRERRKALRAALAAPAIARMPCAFSPLAARAIQEA
EGVTVGAAVADALPDIGITLTVAVHRSROIARVTDLPVLVADATGTGEBMSAAR
TVSELEAGVAGCHLEDOVNRKCGHLDGKGVGTMDIMVERIAAAYNEREDQFVICA
RTDRAVGEIGDSAIRKAYADAGADMIFTEALYSADPEKFAAADVIDPLLANTEFG
KTELLPQALLDEIGINAVIYPTLLIRLAMQGVQALQDANTGIQTDVDRWQHRSL
YELLRYNEYNAFDQOQVFTYGSADSKYKIF"
5658..6809
/gene="Cg10659"

```

```

CDS
5658..6809
/gene="Cg10659"
/EC_number="4.1.3.7"
/notes="PF00365:Citrate synthase"
/codon_start=1
/transl_table=11
/product="Citrate synthase"
/protein_id="BAB98052.1"
/db_xref="GI:21323425"
/translation="MSSATTQVRKGLYGVADYTAVKVMPETNSLTIRGVYVEDLV
ENCSPEVFPVLLMHGELPTAQOLAEFNEGRSVSLDAGLISLIHSLPKEAHMDVMR
TAVSYMGTKDSEYFTTDSHIRKVKHTLLAQPLVLAIRRRKGLDIIAPSSKSA
ENLLSVFGGSPESPASNADVDFEKSLLIYAEHSFNASTFTPARVITTSKSDVYSAI
TGAIGALGPGHGCANEFVMTLAIIDDPKAAANNALDNKVVMPGFGHRTVKGAD
SRVSEKSPRELAAHDGEEKVAMYENRDMADARTGKPNLDPFAGFAYHLLGFPV
DFTPLFVIARVAGWTAHIVEQVNNSLRPLSEYNGEBQREVAPIEK"
complement(6846..7040)
/gene="Cg10660"
CDS
complement(6846..7040)
/gene="Cg10660"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB98053.1"
/db_xref="GI:21323426"
/translation="MYFDISQDYDDELHFSFFELYADQALDAHRCAPHYVWREAVA
KHVLPGSKMVGTRMLHG"
complement(7129..8172)
/gene="Cg10661"
CDS
complement(7129..8172)
/gene="Cg10661"
/EC_number="1.1.1.82"
/notes="PF02615:Malate/L-lactate dehydrogenase"
/codon_start=1
/transl_table=11
/product="Malate/L-lactate dehydrogenases"
/protein_id="BAB98054.1"
/db_xref="GI:21323427"
/translation="MPEVTVNAQQLTVLCTDILTGTVPADAHVLGDSLVQADLMGH
PSHGLRLPMYVRRLHSGAMTHAHEVNLNDLGAVLADGHNGIGOVLAHARKEVT
RAMFGAGVSVNSHFGTAMVYTKAAAGCVSLITTNASAPMAWPGREKRGYN
FWSIAFGETATVVDIANTAVAGKIYHARQTMPIPETWATISGAPTFDPAEIN
GVLPFMAGHKYALISFMMDSVLTGSHSTKHGPDFTFPDGAAGHFLIALDVAF
RDPQDFDALSILVGEYKSTPKAQTETIYPGGESEDRHRKNSAHGISLPEKTMWEL
QELAIENHVVTHR"
complement(8265..8987)
/gene="Cg10662"
CDS
complement(8265..8987)
/gene="Cg10662"
/notes="PF00392:Bacterial regulatory proteins, gntR family"
/codon_start=1
/transl_table=11
/product="Transcriptional regulators"
/protein_id="BAB98055.1"
/db_xref="GI:21323428"
/translation="WVKDAQSPATKADQVYTOIRREIEDGTLPNQMSVWLVE
HTGASRTPVRDARLARLADELIIPEKQAFWPSLSLRHKDLFEPRRIIVEVALEFI
SVGASKSPRIFFGFSFLAADFRELENSADDADFTDRLTSPDLDLVANENHQLFI
RSILSLKPHITTLRIIAHSDHARLRQSVQEHIECEAVASGDLRSAGACRQHLHVE
KSILTLINADSTGSGQIDIRS"
complement(9094..9798)
/gene="Cg10663"
CDS
complement(9094..9798)
/gene="Cg10663"
/codon_start=1
/transl_table=11
/product="Hypothetical membrane protein"
/protein_id="BAB98056.1"
/db_xref="GI:21323429"
/translation="WRFSRLPALITTAVSIPITASATLTADTDKELCTASNTDSSA
VVTWNSEDSVREQRDELDAQPGIKAAIESYIAQDDNAPRAELQVRLDAESGE
GLAMLLPDPFTLADFNABESPTEYTDKAEIISGFSFASDVLSQLQAATIGR
TAETRAEVFADRTDDYNESQALKEDFQNCIDAIDDDARPILPQYILIGGALAVIVL

```

Gene	GIRAWNSRKQSKHSQ"			
CDS	complement(9878..12094) /gene="Cg10664" complement(9878..12094) /gene="Cg10664" /EC number="1.1.1.42" /note="TIGR00178:monomer_idh: isocitrate dehydrogenase, NADP-dependent" /codon_start=1 /transl_table=11 /product="Monomeric isocitrate dehydrogenase (NADP+)" /protein_id="BAB98057.1" /db_xref="GI:21323430" /translation="NAKLIWRTDEAPILATVSLKPVVEAFAATAGIEVETRDISLAG RIIAQPERLTEDQKGNALAEI LELAKTPEANI I KLPNISASVLPQLKAIAKLELODQG YDIPELPDNTTDESKDILARYNAVGSANVPLREGNSDRRAFI AVNFVKFPPHRR			
Query Match	100.0%;	Score 3621;	DB 1;	Length 332050;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3621;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TCGGGCGGGTTAGATCTCGGGGGTTTATTTCACTTTCCTTGGCTGAAGTCGTGACG	60	
DB	37063	TGGGGCGGGTTAGATCTCGGGGGTTTATTTCACTTTCCTTGGCTGAAGTCGTGACG	37122	
QY	61	TCAGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAAACCGATGTTTGAATGGGG	120	
DB	37123	TCAGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAAACCGATGTTTGAATGGGG	37182	
QY	121	AATCGGGGTTACGATACTAGACGCGAGTGACTGCTATCACCTTGGCGGTCTCTTGTG	180	
DB	37183	AATCGGGGTTACGATACTAGACGCGAGTGACTGCTATCACCTTGGCGGTCTCTTGTG	37242	
QY	181	AAAGGAATAATTAATCTAGTGTGCACTCACACATCTTCAAGCTTCCAGCATTCAAAAG	240	
DB	37243	AAAGGAATAATTAATCTAGTGTGCACTCACACATCTTCAAGCTTCCAGCATTCAAAAG	37302	
QY	241	ATCTTGTGAGCAACCGCGCGAAATTCGGGTCCTGCTTTCCTGTCGAGCACTCGAAACC	300	
DB	37303	ATCTTGTGAGCAACCGCGCGAAATTCGGGTCCTGCTTTCCTGTCGAGCACTCGAAACC	37362	
QY	301	GGTGAGCAGCAGGTAGCTATTATCCCGCTGAAGATCGGGATCATTCACCGCTCTTTT	360	
DB	37363	GGTGAGCAGCAGGTAGCTATTATCCCGCTGAAGATCGGGATCATTCACCGCTCTTTT	37422	
QY	361	GCTTCTGAAGCTGTCGCAATTTGTATCCGAAGGCTCACAGTCAAGCGGTACTGGAATC	420	
DB	37423	GCTTCTGAAGCTGTCGCAATTTGTATCCGAAGGCTCACAGTCAAGCGGTACTGGAATC	37482	
QY	421	GATGAAATTAATCGGTGACGTAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC	480	
DB	37483	GATGAAATTAATCGGTGACGTAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC	37542	
QY	481	TTCTCTGTGAAATGCCAGCTTTCGCCGAGTGTGCGGAAACGGCATTTACTTTTATT	540	
DB	37543	TTCTCTGTGAAATGCCAGCTTTCGCCGAGTGTGCGGAAACGGCATTTACTTTTATT	37602	
QY	541	GGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAAACCGCGCG	600	
DB	37603	GGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAAACCGCGCG	37662	
QY	601	AAGAAGGCTGTCTGCCAGTTTTCGGCGAATCCACCCGAGCAAAAACATCGATGAGATC	660	
DB	37663	AAGAAGGCTGTCTGCCAGTTTTCGGCGAATCCACCCGAGCAAAAACATCGATGAGATC	37722	
QY	661	GTTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGAGTGTCCGGTGGTGC	720	
DB	37723	GTTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGAGTGTCCGGTGGTGC	37782	
QY	721	GGAGCGGTTAGCGTTTGTGTTCTCACTGATGAGCTTCGCAAAATAGCAACAGAGCA	780	
DB	37783	GGAGCGGTTAGCGTTTGTGTTCTCACTGATGAGCTTCGCAAAATAGCAACAGAGCA	37842	

Qy	781	TCTCGTGAAGCTGAAGCGGCTTTTCGGGATGCGCGGTATATGTGCAACGTCGTGTGATT 840		
Db	37843	TCTCGTGAAGCTGAAGCGGCTTTTCGGGATGCGCGGTATATGTGCAACGTCGTGTGATT 37902		
Qy	841	AACCTCAGCATATTGAAGTCAGATCTTGGCCATCACACTGAGGAAGTTGTACACCTT 900		
Db	37903	AACCTCAGCATATTGAAGTCAGATCTTGGCCATCACACTGAGGAAGTTGTACACCTT 37962		
Qy	901	TATGAACGTGACTGCTCACTGACGCTGTCACCAAAAAGTTGTGCAAAATTCGCCACGA 960		
Db	37963	TATGAACGTGACTGCTCACTGACGCTGTCACCAAAAAGTTGTGCAAAATTCGCCACGA 38022		
Qy	961	CAGCATTTGGATCCAGAACTCGGTGATCGCATTTTGGGATGCACTAAAGTTCTGGCG 1020		
Db	38023	CAGCATTTGGATCCAGAACTCGGTGATCGCATTTTGGGATGCACTAAAGTTCTGGCG 38082		
Qy	1021	TCCATTGGTTACCAAGGCGCGGAAACCGTGGAAATTCCTTGGTCGATGAAAAGGCAACAC 1080		
Db	38083	TCCATTGGTTACCAAGGCGCGGAAACCGTGGAAATTCCTTGGTCGATGAAAAGGCAACAC 38142		
Qy	1081	GTCTTCATCGAAATGAACCCAGATTCAGAGTTGAGCAACCCGTGACTGAAGAAGTCACC 1140		
Db	38143	GTCTTCATCGAAATGAACCCAGATTCAGAGTTGAGCAACCCGTGACTGAAGAAGTCACC 38202		
Qy	1141	GAGGTGGAACCTGGTGAAGGCGCAGATCGCTTGGCTGCTGGTGCACCTTCAAGGAATTG 1200		
Db	38203	GAGGTGGAACCTGGTGAAGGCGCAGATCGCTTGGCTGCTGGTGCACCTTCAAGGAATTG 38262		
Qy	1201	GGTTCGACCCAGATAAGATCAAGACCCAGTCAGTGACGCTGCGGATGCGCATCACAG 1260		
Db	38263	GGTTCGACCCAGATAAGATCAAGACCCAGTCAGTGACGCTGCGGATGCGCATCACAG 38322		
Qy	1261	GAAGATCAAAACACCGCTTCCGCCAGATACCGAACTATCACCGCTACCGTCAACA 1320		
Db	38323	GAAGATCAAAACACCGCTTCCGCCAGATACCGAACTATCACCGCTACCGTCAACA 38382		
Qy	1321	GGCGAGCTGGCTGCTTGAAGGTGCACTCAGCTCGGTGGGGAATCACCGCAAC 1380		
Db	38383	GGCGAGCTGGCTGCTTGAAGGTGCACTCAGCTCGGTGGGGAATCACCGCAAC 38442		
Qy	1381	TTTGACTCCATGCTGCTCAAAATGACCTCGCTGGTTCGCACTTTGAAACTGCTGTCT 1440		
Db	38443	TTTGACTCCATGCTGCTCAAAATGACCTCGCTGGTTCGCACTTTGAAACTGCTGTCT 38502		
Qy	1441	CGTGACAGCGCGCTTGGCTGAGTTACCGTGTCTGCTGTGCAACCAACATTTGTTTC 1500		
Db	38503	CGTGACAGCGCGCTTGGCTGAGTTACCGTGTCTGCTGTGCAACCAACATTTGTTTC 38562		
Qy	1501	TTGCGTGGTGGTGGGAGAGGACTTCACTTCCAGCGCATCGCCACCGGATTCATT 1560		
Db	38563	TTGCGTGGTGGTGGGAGAGGACTTCACTTCCAGCGCATCGCCACCGGATTCATT 38622		
Qy	1561	GCCGATCACCCGCACTCTCTTTCAGGCTCCACTGCTGATGATGAGCAGGAGCGCATCTG 1620		
Db	38623	GCCGATCACCCGCACTCTCTTTCAGGCTCCACTGCTGATGATGAGCAGGAGCGCATCTG 38682		
Qy	1621	GATTAATGCGCAGATGTCACCGTGAAGGACTTCACTTCCAGCGCATCGCCACCGGATTCGA 1680		
Db	38683	GATTAATGCGCAGATGTCACCGTGAAGGACTTCACTTCCAGCGCATCGCCACCGGATTCGA 38742		
Qy	1681	GCTCTCTCGATAAGCTGCTAAACATCAAGGATCTGCCACTGCCACCGGTTCCCGTGAC 1740		
Db	38743	GCTCTCTCGATAAGCTGCTAAACATCAAGGATCTGCCACTGCCACCGGTTCCCGTGAC 38802		
Qy	1741	CGCTGAGCAGCTTGGCCGAGCGGTTGCTGCTGATCTCCGTCGAGCAGGACGCACTG 1800		
Db	38803	CGCTGAGCAGCTTGGCCGAGCGGTTGCTGCTGATCTCCGTCGAGCAGGACGCACTG 38862		
Qy	1801	CGAGTTACTGTATACCACTTCCGCGATGCACACCACTCTTTTGTGCGACCGAGTCGCG 1860		
Db	38863	CGAGTTACTGTATACCACTTCCGCGATGCACACCACTCTTTTGTGCGACCGAGTCGCG 38922		
Qy	1861	TCAATTCGCACTGAAGCTGCGGCGAGGCGGTGCGCAAAAGCTGACTCTCTGAGCTTTTGTTC 1920		

Db 38923 TCATTGCACTGAAGCTTCGGCAGAGCGCTCGCAAGCTGACTCCTGAGCTTTGTGCC 38982  
Qy 1921 GTGAGGCTTGGGGCGGCGGACCTACGATGTGGGATGCGTTTCTCTTTGAGATCGG 1980  
Db 38983 GTGAGGCTTGGGGCGGCGGACCTACGATGTGGGATGCGTTTCTCTTTGAGATCGG 39042  
Qy 1981 TGGGACAGGCTCGACAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040  
Db 39043 TGGGACAGGCTCGACAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 39102  
Qy 2041 CGGGCGGCAACACCTGGGATACACCCGTCACCGATCCGATCCGTCGCGCGCTTGT 2100  
Db 39103 CGGGCGGCAACACCTGGGATACACCCGTCACCGATCCGATCCGTCGCGCGCTTGT 39162  
Qy 2101 AAGGAAGCTGCCAGCTCCGCGCTGGACATCTTCGCACTTCGACGCGCTTAAACAGCTC 2160  
Db 39163 AAGGAAGCTGCCAGCTCCGCGCTGGACATCTTCGCACTTCGACGCGCTTAAACAGCTC 39222  
Qy 2161 TCCAGATGGTCCAGATCCGAGCGAGTCCGTCGAGACCAACACCGCGTAGCGAGGTG 2220  
Db 39223 TCCAGATGGTCCAGATCCGAGCGAGTCCGTCGAGACCAACACCGCGTAGCGAGGTG 39282  
Qy 2221 GCTATGCTTATCTGCTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCCTGGATTAC 2280  
Db 39283 GCTATGCTTATCTGCTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCCTGGATTAC 39342  
Qy 2281 TACCTAAGATGGCAGAGGATCTGTCAGTCTGGCGCTCACATCTTGGCCATTAGGAT 2340  
Db 39343 TACCTAAGATGGCAGAGGATCTGTCAGTCTGGCGCTCACATCTTGGCCATTAGGAT 39402  
Qy 2341 ATGGCTGCTGCTTCCGCCAGCTGGGTAAACCAAGCTGTCACCGCACTGGCGCGTGAA 2400  
Db 39403 ATGGCTGCTGCTTCCGCCAGCTGGGTAAACCAAGCTGTCACCGCACTGGCGCGTGAA 39462  
Qy 2401 TTGATCTGCCAGTGCAAGTGCAACCCAGACACTGCGGCTGGCCAGCTGGCAACCTAC 2460  
Db 39463 TTGATCTGCCAGTGCAAGTGCAACCCAGACACTGCGGCTGGCCAGCTGGCAACCTAC 39522  
Qy 2461 TTTGCTCAGCTCAAGCTGTGAGATGCTGTTGAGCGGTGCTTCGCGACCACTGTCTGGC 2520  
Db 39523 TTTGCTCAGCTCAAGCTGTGAGATGCTGTTGAGCGGTGCTTCGCGACCACTGTCTGGC 39582  
Qy 2521 ACCACTCCAGGCAATCCCTGTGCAATGTTGCTGCAATGCGCACACCGCTGCGGAT 2580  
Db 39583 ACCACTCCAGGCAATCCCTGTGCAATGTTGCTGCAATGCGCACACCGCTGCGGAT 39642  
Qy 2581 ACCGTTTGGAGCTCGAGGCTGTTTCTGACTCGAGCGTACTGGAAGCAGTGGCGGA 2640  
Db 39643 ACCGTTTGGAGCTCGAGGCTGTTTCTGACTCGAGCGTACTGGAAGCAGTGGCGGA 39702  
Qy 2641 CTGTACTGCAATTTGATCTGGAAACCCAGGCGCCAAACCGGTGCGGTCTACCGCCACGAA 2700  
Db 39703 CTGTACTGCAATTTGATCTGGAAACCCAGGCGCCAAACCGGTGCGGTCTACCGCCACGAA 39762  
Qy 2701 ATCCAGCGGACAGTTGTCACCTCGGTGCAAGCGCCACCGCTGCGGCTTGGCGAT 2760  
Db 39763 ATCCAGCGGACAGTTGTCACCTCGGTGCAAGCGCCACCGCTGCGGCTTGGCGAT 39822  
Qy 2761 CGTTTCGAATCATCGAAGCAACTACGACGCGGTTAATGAGATGCTGGGACGCCCAACC 2820  
Db 39823 CGTTTCGAATCATCGAAGCAACTACGACGCGGTTAATGAGATGCTGGGACGCCCAACC 39882  
Qy 2821 AAGTCAACCCATCTCCAGTTGTTGGGACCTCGCACTCCACCTCTGTTGGTGGGCT 2880  
Db 39883 AAGTCAACCCATCTCCAGTTGTTGGGACCTCGCACTCCACCTCTGTTGGTGGGCT 39942  
Qy 2881 GTGATCCAGCAGACTTTGTCTGCCGATCCCAAAAGTACGACATCCAGACTCTGTATC 2940  
Db 39943 GTGATCCAGCAGACTTTGTCTGCCGATCCCAAAAGTACGACATCCAGACTCTGTATC 40002  
Qy 2941 CGTTCTCGCGGCGAGCTTGGTAAACCTCCAGGTGCTGGCGAGCCACTGGCGACC 3000

Db 40003 GCGTCTCTCGCGCGAGCTTGGTAAACCTCCAGGTGGTGGCCAGAGCCACTGGCGACC 40062  
Qy 3001 CGCGCACTGGAAGGCGCTCCGAAAGCAAGGCACTCTGACGGAAGTTCTCTGAGGAAGAG 3060  
Db 40063 CGCGCACTGGAAGGCGCTCCGAAAGCAAGGCACTCTGACGGAAGTTCTCTGAGGAAGAG 40122  
Qy 3061 CAGCGCACTCGACGCTGATGTTCCAAAGCAAGTTCGCAATAGCTCAACGCGCTGCTG 3120  
Db 40123 CAGCGCACTCGACGCTGATGTTCCAAAGCAAGTTCGCAATAGCTCAACGCGCTGCTG 40182  
Qy 3121 TTCCGAAGCAACCAAGAGTTCTCTCGAGCACCGTCCGCGCTTCGSCAACACCTCTGCG 3180  
Db 40183 TTCCGAAGCAACCAAGAGTTCTCTCGAGCACCGTCCGCGCTTCGSCAACACCTCTGCG 40242  
Qy 3181 CTGATGATCGTAATCTTCTACGCGCTGGTTCGAAGGCGGAGACTTTGATCCGCGCTG 3240  
Db 40243 CTGATGATCGTAATCTTCTACGCGCTGGTTCGAAGGCGGAGACTTTGATCCGCGCTG 40302  
Qy 3241 CCAGATGCGGCAACCCACCTGCTTGTTCGCTCGATCGATCTCTGAGCCAGACGATAAG 3300  
Db 40303 CCAGATGCGGCAACCCACCTGCTTGTTCGCTCGATCGATCTCTGAGCCAGACGATAAG 40362  
Qy 3301 GGTATGCGCAATGTTGTGGCAACGTCACCGCCAGATCCGCGCAATGCGTGGCGTAC 3360  
Db 40363 GGTATGCGCAATGTTGTGGCAACGTCACCGCCAGATCCGCGCAATGCGTGGCGTAC 40422  
Qy 3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAAGAAAGCAGATTCTCTCAACAAAGGCGCAT 3420  
Db 40423 CGCTCCGTTGAGTCTGTCACCGCAACCGCAAGAAAGCAGATTCTCTCAACAAAGGCGCAT 40482  
Qy 3421 GTTCTCGACATTCGCTGCTGTTGTCACCGTCACTGCTGCTGCTGCTGCTGCTGCTGCT 3480  
Db 40483 GTTCTCGACATTCGCTGCTGTTGTACCGTCACTGCTGCTGCTGCTGCTGCTGCTGCT 40542  
Qy 3481 GCTGAGATGAGTTCGCAATCATCGAGGCTATGAAGTGAAGCAACAATCACTGCTTCT 3540  
Db 40543 GCTGAGATGAGTTCGCAATCATCGAGGCTATGAAGTGAAGCAACAATCACTGCTTCT 40602  
Qy 3541 GTTCAAGCAAAATGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600  
Db 40603 GTTCAAGCAAAATGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 40662  
Qy 3601 TTGATCGCTGCTGCTTCTTAA 3621  
Db 40663 TTGATCGCTGCTGCTTCTTAA 40683

## RESULT 5

AX127145 34980 bp DNA linear PAT 11-MAY-2001  
LOCUS Sequence 7061 from Patent EP1108790.

## DEFINITION

AX127145 AX114121

## ACCESSION

AX127145.1 GI:14041133

## VERSION

KEYWORDS

## SOURCE

ORGANISM

Corynebacterium glutamicum

Corynebacterium glutamicum

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Corynebacterium.

## REFERENCE

1

AUTHORS

Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,

Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.

Novel polynucleotides

Patent: EP 1108790-A 7061 20-JUN-2001;

KYOWA HAKKO KOGYO CO., LTD. (JP)

LOCATION/Qualifiers

1..34980

/organism="Corynebacterium glutamicum"

/mol\_type="genomic DNA"

/db\_xref="taxon:1718"

/note="Seq 1 to long (3,309,400) split in 11, seq 7061

0,600,001,0,949,980,"

## ORIGIN



Query Match				100.0%; Score 3621; DB 6; Length 349980;			
Best Local Similarity				100.0%; Pred. No. 0;			
Matches 3621; Conservative				0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	TGGGGCGGGTTAGATCCCTGGGGGGTTTATTTCACTTTGGCTTGAAGTGTGTGAGG	60				
Db	105013	TGGGGCGGGTTAGATCCCTGGGGGGTTTATTTCACTTTGGCTTGAAGTGTGTGAGG	105072				
QY	61	TCAGGGGAGTGTGGCCGAAACATTGAGAGGAAACAAACACGATGTTGATTGGGGG	120				
Db	105073	TCAGGGGAGTGTGGCCGAAACATTGAGAGGAAACAAACACGATGTTGATTGGGGG	105132				
QY	121	AATCGGGGTTACGATATAGAGCGAGTGAATCGCGTCCGTTCCGTCAGCTTGTGTTG	180				
Db	105133	AATCGGGGTTACGATATAGAGCGAGTGAATCGCGTCCGTTCCGTCAGCTTGTGTTG	105192				
QY	181	AAAGGAATAATTACTCTAGTGTGCACTCACATCTTCAAGCGTTCCAGCATTCGAAAG	240				
Db	105193	AAAGGAATAATTACTCTAGTGTGCACTCACATCTTCAAGCGTTCCAGCATTCGAAAG	105252				
QY	241	ATCTTGGTAGCAACCGCGGGAATCGCGTCCGTTCCGTCAGCACTCGAAACC	300				
Db	105253	ATCTTGGTAGCAACCGCGGGAATCGCGTCCGTTCCGTCAGCACTCGAAACC	105312				
QY	301	GGTCAGCGCAGGTAGCTATTACCCCGTGAAGATCGGGATCATTTCCACCGCTCTTTT	360				
Db	105313	GGTCAGCGCAGGTAGCTATTACCCCGTGAAGATCGGGATCATTTCCACCGCTCTTTT	105372				
QY	361	GCTTCTGAAGCTGTCCGCAATTTGGTACCGAAGCTCACCAAGCGTACCTGGACATC	420				
Db	105373	GCTTCTGAAGCTGTCCGCAATTTGGTACCGAAGCTCACCAAGCGTACCTGGACATC	105432				
QY	421	GATGAATAATTCGGTGCAGCTTAAAGATTAAGCAGATGCCATTTACCCGGATACGGC	480				
Db	105433	GATGAATAATTCGGTGCAGCTTAAAGATTAAGCAGATGCCATTTACCCGGATACGGC	105492				
QY	481	TTCTCTGTCTGAAATATGCCAGCTTGCCTCCGAGTGTGCGGAAACGGCATTTATT	540				
Db	105493	TTCTCTGTCTGAAATATGCCAGCTTGCCTCCGAGTGTGCGGAAACGGCATTTATT	105552				
QY	541	GGCCCAACCCGAGAGTTCCTGATCTCACCGTGTATAGTCTCGCGGTAAACCGCGG	600				
Db	105553	GGCCCAACCCGAGAGTTCCTGATCTCACCGTGTATAGTCTCGCGGTAAACCGCGG	105612				
QY	601	AAGAGGCTGTCTCCAGTTTGGCGGAATCCACCCGAGCAAAACATCGATGAGATC	660				
Db	105613	AAGAGGCTGTCTCCAGTTTGGCGGAATCCACCCGAGCAAAACATCGATGAGATC	105672				
QY	661	GTAAAGCGCTGAAGGCGCAGACTTACCCCATTTTGTGAAGGCGAGTTGCCGTGGC	720				
Db	105673	GTAAAGCGCTGAAGGCGCAGACTTACCCCATTTTGTGAAGGCGAGTTGCCGTGGC	105732				
QY	721	GGACCGGTATCGTTTTTGTGCTTCACTGATGAGCTTCGCAATAGCAACAGAGCA	780				
Db	105733	GGACCGGTATCGTTTTTGTGCTTCACTGATGAGCTTCGCAATAGCAACAGAGCA	105792				
QY	781	TCTCGTGAAGCTGAAGCGCTTTCGCGATGCGCGGTATATGCGAAGCTGTGATT	840				
Db	105793	TCTCGTGAAGCTGAAGCGCTTTCGCGATGCGCGGTATATGCGAAGCTGTGATT	105852				
QY	841	AACCTCTCAGCATATTGAGTGCAGATCCCTGCGCATCACTGGAGAGTTGTACACCTT	900				
Db	105853	AACCTCTCAGCATATTGAGTGCAGATCCCTGCGCATCACTGGAGAGTTGTACACCTT	105912				
QY	901	TATGAACGTGACTGTCTACCTGAGCGTGTCTACCAAAAGTTGTGAAATTTGCGCCAGCA	960				
Db	105913	TATGAACGTGACTGTCTACCTGAGCGTGTCTACCAAAAGTTGTGAAATTTGCGCCAGCA	105972				
QY	961	CAGCATTTGGATCCAGATCGGTGATCGCATTTTGTGGGATGCAAGTTCCTGCGG	1020				
Db	105973	CAGCATTTGGATCCAGATCGGTGATCGCATTTTGTGGGATGCAAGTTCCTGCGG	106032				
QY	1021	TCCATTGGTTACGAGCGCGGGAACCGGTGGAATTTCTTGGTCGATGAAAGGCGCAACAC	1080				
Db	106033	TCCATTGGTTACGAGCGCGGGAACCGGTGGAATTTCTTGGTCGATGAAAGGCGCAACAC	106092				
QY	1081	GTCTTCATCGGAATGAACCCAGTATCCAGTTTGAGCACACCGTGAAGAGTCAACC	1140				
Db	106093	GTCTTCATCGGAATGAACCCAGTATCCAGTTTGAGCACACCGTGAAGAGTCAACC	106152				
QY	1141	GAGTGGACCTGGTGAAGCGCGAGATGCGCTTGGCTGTGTGTGCAACCTTGAAGGAATTG	1200				
Db	106153	GAGTGGACCTGGTGAAGCGCGAGATGCGCTTGGCTGTGTGTGCAACCTTGAAGGAATTG	106212				
QY	1201	GGTCTGACCCAGATAAGATCAAGACCCAGTGCAGCTGCAGTGCAGGATCAACACG	1260				
Db	106213	GGTCTGACCCAGATAAGATCAAGACCCAGTGCAGCTGCAGTGCAGGATCAACACG	106272				
QY	1261	GAGATCCAAAACAGCGCTTCCGCCAGATACCGGAATATCACCGCTACCGCTCACCA	1320				
Db	106273	GAGATCCAAAACAGCGCTTCCGCCAGATACCGGAATATCACCGCTACCGCTCACCA	106332				
QY	1321	GGCGAGCTGGGCTTGGCTTGAACGCTGAGCTGAGCTGCGTGGGGAATCACCGCAC	1380				
Db	106333	GGCGAGCTGGGCTTGGCTTGAACGCTGAGCTGAGCTGCGTGGGGAATCACCGCAC	106392				
QY	1381	TTTGACTCCATGCTGGTGAATAATGACCTGCCGCTGCTCCGACTTTGAAACTGCTGTGT	1440				
Db	106393	TTTGACTCCATGCTGGTGAATAATGACCTGCCGCTGCTCCGACTTTGAAACTGCTGTGT	106452				
QY	1441	CGTGACAGCGCGTGGCTGAGTTACCGTGTCTGCTGTGTGTGCAACCAATGCTGTTTC	1500				
Db	106453	CGTGACAGCGCGTGGCTGAGTTACCGTGTCTGCTGTGTGTGCAACCAATGCTGTTTC	106512				
QY	1501	TTGCGTGGCTTGTGCGGGAAGAGCACTTCACTTCCAGCGCATCGCCACCGGATTCATT	1560				
Db	106513	TTGCGTGGCTTGTGCGGGAAGAGCACTTCACTTCCAGCGCATCGCCACCGGATTCATT	106572				
QY	1561	GGCGATCAACCGCACTCTCTTCAAGCTTCACTGCTGATGATGAGGAGGAGCATCTCTG	1620				
Db	106573	GGCGATCAACCGCACTCTCTTCAAGCTTCACTGCTGATGATGAGGAGGAGCATCTCTG	106632				
QY	1621	GATTACTTGGCAGATGTACCGTGAACCAAGCTCATGTTGCTGCTCCAAAGATGTTGCA	1680				
Db	106633	GATTACTTGGCAGATGTACCGTGAACCAAGCTCATGTTGCTGCTCCAAAGATGTTGCA	106692				
QY	1681	GCTCTTATCGATAAGCTGCTTAAACATCAAGATCTGCCACTGCCACGGGTTCCCGTGAC	1740				
Db	106693	GCTCTTATCGATAAGCTGCTTAAACATCAAGATCTGCCACTGCCACGGGTTCCCGTGAC	106752				
QY	1741	CGCTGAAGCAGCTTGGCGCAGCGGTTTGTCTGCTGATCTCCGTGAGCAGGACGCACTG	1800				
Db	106753	CGCTGAAGCAGCTTGGCGCAGCGGTTTGTCTGCTGATCTCCGTGAGCAGGACGCACTG	106812				
QY	1801	GCAATTACTGATACCACTTCCCGCATGCAACCACTTTCCTTTCGACCCGAGTCCG	1860				
Db	106813	GCAATTACTGATACCACTTCCCGCATGCAACCACTTTCCTTTCGACCCGAGTCCG	106872				
QY	1861	TGATTCGCACTGAGCCTTGGCGAGGCGCTGCAAGCTGCACTCTGAGCTTTTGTCC	1920				
Db	106873	TGATTCGCACTGAGCCTTGGCGAGGCGCTGCAAGCTGCACTCTGAGCTTTTGTCC	106932				
QY	1921	GTGAGGCTTGGGGCGCGCACTTACGATGTGGCGATCGCTTTCCTTTTGAAGATCCG	1980				
Db	106933	GTGAGGCTTGGGGCGCGCACTTACGATGTGGCGATCGCTTTCCTTTTGAAGATCCG	106992				
QY	1981	TGGGACAGGCTGACAGCTGCGGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT	2040				
Db	106993	TGGGACAGGCTGACAGCTGCGGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT	107052				
QY	2041	CGCGGCGCAACACCGTGGGATACACCCGTACCCAGACTCCGCTGTGCGCGGCTTGT	2100				
Db	107053	CGCGGCGCAACACCGTGGGATACACCCGTACCCAGACTCCGCTGTGCGCGGCTTGT	107112				
QY	2101	AAGGAAGCTGCGAGCTTCCGGGTGAGCATCTTCCGATCTTTCGACCGGCTTAAACGCTC	2160				
Db	107113	AAGGAAGCTGCGAGCTTCCGGGTGAGCATCTTCCGATCTTTCGACCGGCTTAAACGCTC	107172				



Db 107113 AAGGAAGCTGCCAGCTCCGGCGTGACATCTTCGCCATCTTCGACGCCCTTAAACGACGTC 107172  
Qy 2161 TCCAGATGCGTCGAGCAATGACGCGAGTCTTGAGAGCAACACCGCGGTAGCGAGGTG 2220  
Db 107173 TCCAGATGCGTCGAGCAATGACGCGAGTCTTGAGAGCAACACCGCGGTAGCGAGGTG 107232  
Qy 2221 GCTATGGCTTATCTTGCTGATCTCTGATCCAAATGAAAGCTCTACACCTGATATAC 2280  
Db 107233 GCTATGGCTTATCTTGCTGATCTCTGATCCAAATGAAAGCTCTACACCTGATATAC 107292  
Qy 2281 TACCTAAAGATGGCAGAGAGATGCTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340  
Db 107293 TACCTAAAGATGGCAGAGAGATGCTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 107352  
Qy 2341 ATGECTGGCTCTGCTTCGCCAGCTGGCGTAAACAAAGCTGGTCAACCGCACTGGCGCGTAA 2400  
Db 107353 ATGECTGGCTCTGCTTCGCCAGCTGGCGTAAACAAAGCTGGTCAACCGCACTGGCGCGTAA 107412  
Qy 2401 TTGATCTGCCAGTGCATGTCACACCCACGACATCTGGCGGTGGCAGTGGCAACCTTAC 2460  
Db 107413 TTGATCTGCCAGTGCATGTCACACCCACGACATCTGGCGGTGGCAGTGGCAACCTTAC 107472  
Qy 2461 TTGCTGACAGCTCAAGCTGTCAGATGCTGTGACGGTGTTCGCGACCACTGTCTGGC 2520  
Db 107473 TTGCTGACAGCTCAAGCTGTCAGATGCTGTGACGGTGTTCGCGACCACTGTCTGGC 107532  
Qy 2521 ACCACTCCAGCACTCCCTGTCTGCAATGTTGTGTCATTCGCGACACACCGTGGCGAT 2580  
Db 107533 ACCACTCCAGCACTCCCTGTCTGCAATGTTGTGTCATTCGCGACACACCGTGGCGAT 107592  
Qy 2581 ACCGGTTGAGCTCGAGCTGTTCTGACCTCGAGCGTACTGCGAAGCAGTGGCGGA 2640  
Db 107593 ACCGGTTGAGCTCGAGCTGTTCTGACCTCGAGCGTACTGCGAAGCAGTGGCGGA 107652  
Qy 2641 CTGTACTGCACTTTCAGTCTGGAACCCAGCGCCCAACCGGTGCGCTTACCGCCACGAA 2700  
Db 107653 CTGTACTGCACTTTCAGTCTGGAACCCAGCGCCCAACCGGTGCGCTTACCGCCACGAA 107712  
Qy 2701 ATCCAGCGGACAGTGTCCACCTGCTGACAGCGCCACCGCACTGGCGCTTGGGAT 2760  
Db 107713 ATCCAGCGGACAGTGTCCACCTGCTGACAGCGCCACCGCACTGGCGCTTGGGAT 107772  
Qy 2761 CGTTTCGAATCATCGAAGCAACTACGACGCGGTTAATGAGATGCTGGGACGCCCAACC 2820  
Db 107773 CGTTTCGAATCATCGAAGCAACTACGACGCGGTTAATGAGATGCTGGGACGCCCAACC 107832  
Qy 2821 AAGTCAACCCATCTCCAGGTTGTCGAGCTCGCACTCCACTCGCTTGGTGGCGGT 2880  
Db 107833 AAGTCAACCCATCTCCAGGTTGTCGAGCTCGCACTCCACTCGCTTGGTGGCGGT 107892  
Qy 2881 GTGGATCCAGCAGACTTGTGCGCGATCCACAAAGTACGACATCCAGACTCTGTATC 2940  
Db 107893 GTGGATCCAGCAGACTTGTGCGCGATCCACAAAGTACGACATCCAGACTCTGTATC 107952  
Qy 2941 GCGTTCCTGCGCGGAGCTGTGTAACCTTCAGGTGGTGGCCAGAGCACTGGCCACC 3000  
Db 107953 GCGTTCCTGCGCGGAGCTGTGTAACCTTCAGGTGGTGGCCAGAGCACTGGCCACC 108012  
Qy 3001 CGCGCATGGAAGCGCGCTCCGAGGCAAGGACCTCTGACGGAAGTTCCTGAGGAAGAG 3060  
Db 108013 CGCGCATGGAAGCGCGCTCCGAGGCAAGGACCTCTGACGGAAGTTCCTGAGGAAGAG 108072  
Qy 3061 CAGCGCACTCGACGCTGATGATTCGAGGAAGCTGCGAATGACCTCAACCGCTGCTG 3120  
Db 108073 CAGCGCACTCGACGCTGATGATTCGAGGAAGCTGCGAATGACCTCAACCGCTGCTG 108132  
Qy 3121 TTCCGGAAGCAACCGGAAGTTCCTCGAGCAACGCTGCGCGCTTCGGCAACACCTCTGG 3180  
Db 108133 TTCCGGAAGCAACCGGAAGTTCCTCGAGCAACGCTGCGCGCTTCGGCAACACCTCTGG 108192  
Qy 3181 CTGATGATGCTGAATTCCTTACGCGCTGGTCAAGGCGCGAGACTTTGATCCGCGCTG 3240  
Db 108193 CTGATGATGCTGAATTCCTTACGCGCTGGTCAAGGCGCGAGACTTTGATCCGCGCTG 108252

Qy 3241 CCAGATGTCGCGACCCCACTGCTGTTGTCCTGGATGCGATCTCTGAGCCACGATAAG 3300  
Db 108253 CCAGATGTCGCGACCCCACTGCTGTTGTCCTGGATGCGATCTCTGAGCCACGATAAG 108312  
Qy 3301 GGTATCGGCAATGTTGTGTCGCAACGTCACAGCGCCAGATCCGCCCAATGCGTGTGCGTAC 3360  
Db 108313 GGTATCGGCAATGTTGTGTCGCAACGTCACAGCGCCAGATCCGCCCAATGCGTGTGCGTAC 108372  
Qy 3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGAGATTCCTCCAAAGGCGCCAT 3420  
Db 108373 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGAGATTCCTCCAAAGGCGCCAT 108432  
Qy 3421 GTTCTGCAACCAATTCGCTGGTGTGTCACCGTCACTGTTCCTGGAAGTGTGATGAGTCAAG 3480  
Db 108433 GTTCTGCAACCAATTCGCTGGTGTGTCACCGTCACTGTTCCTGGAAGTGTGATGAGTCAAG 108492  
Qy 3481 GCTGGAGATGCACTGCGCAATCATCGAGGTATGAGATGGAAGCAACAATCACTGCTTCT 3540  
Db 108493 GCTGGAGATGCACTGCGCAATCATCGAGGTATGAGATGGAAGCAACAATCACTGCTTCT 108552  
Qy 3541 GTTACCGGCAAAATCGATCGCTGTTGTTGCTGTCGCAACGAGTGGAAAGTGGCGAC 3600  
Db 108553 GTTACCGGCAAAATCGATCGCTGTTGTTGCTGTCGCAACGAGTGGAAAGTGGCGAC 108612  
Qy 3601 TTGATGCTGCTGCTTTCCTAA 3621  
Db 108613 TTGATGCTGCTGCTTTCCTAA 108633

RESULT 6  
LOCUS CGPYC 3728 bp DNA linear BCT 08-MAY-1998  
DEFINITION Corynebacterium glutamicum pyc gene.  
ACCESSION Y05548  
VERSION Y05548.1 GI:2879822  
KEYWORDS pyc gene; pyruvate carboxylase.  
SOURCE Corynebacterium glutamicum  
ORGANISM Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1 Peters-Wendisch, P.G., Kreutzer, C., Kalinowski, J., Patek, M., Sahm, H. and Sikkema, B.J.  
Pyruvate carboxylase from Corynebacterium glutamicum: characterization, expression and inactivation of the pyc gene  
Microbiology (Reading, Engl.) 144 (Pt 4), 915-927 (1998)  
98240228  
PUBMED 9579055

REFERENCE 2 (bases 1 to 3728)  
Peters-Wendisch, P.G.  
Direct Submission  
Submitted (21-NOV-1996) P.G. Peters-Wendisch, Institut fuer  
Biotechnologie 1, Forschungszentrum Juelich GmbH, Juelich, D-52425,  
FRG

FEATURES  
Location/Qualifiers  
1. 3728  
/organism="Corynebacterium glutamicum"  
/mol\_type="genomic DNA"  
/strain="ATCC 13032"  
/db\_xref="taxon:1718"  
109..3587  
/gene="pyc"  
109..3587  
/evidence=experimental  
165..3587  
/gene="pyc"  
/EC\_number="6.4.1.1"  
/codon\_start=1  
/transl\_table=11  
/product="pyruvate carboxylase"  
/protein\_id="CAA70739.1"

44	QY	GCTTGAAGTCGTGCAGGTCAGGGGAGTGTGCCCGAATAACATTTAGAGGGAACACAAAAC	103
10	DB	GCTTGAAGTCGTGCAGGTCAGGGAGTGTGCCCGAATAACATTTAGAGGGAACACAAAAC	69
104	QY	CGATGTTTGTATGGGGGAATCGGGGTTACGATACTAGGACGCACTGCTATCACCC	163
70	DB	CGATGTTTGTATGGGGGAATCGGGGTTACGATACTAGGACGCACTGCTATCACCC	129
164	QY	TTGCGGTCCTTCGTTGAAAGGAATAATTACTCTTAGTGTGCACATCACACATCTTCAACGC	223
130	DB	TTGCGGTCCTTCGTTGAAAGGAATAATTACTCTTAGTGTGCACATCACATCTTCAACGC	189
224	QY	TTCCAGCAATTCAAAAAGATCTTGGTAGCAAAACCGCGGCGAAATCGGGTCGGTCTTCC	283
190	DB	TTCCAGCAATTCAAAAAGATCTTGGTAGCAAAACCGCGGCGAAATCGGGTCGGTCTTCC	249
284	QY	GTGCAGCACTCGAAACCGGTGCAGCCAGCTAGCTATTATCCCGCTGAAAGATCGGGGAT	343
250	DB	GTGCAGCACTCGAAACCGGTGCAGCCAGCTAGCTATTATCCCGCTGAAAGATCGGGGAT	309
344	QY	CATTCCACCGCTCTTTTGCCTCTGAAGCTGCCGATTTGGTACCGAAGGCTCACAGTCA	403
310	DB	CATTCCACCGCTCTTTTGCCTCTGAAGCTGCCGATTTGGTACCGAAGGCTCACAGTCA	369
404	QY	AGGGGTACTCTGCACATCATGAAATTATTCGGTGCAGCTTAAAAAGTTAAAGCAGATGCCA	463
370	DB	AGGGGTACTCTGCACATCATGAAATTATTCGGTGCAGCTTAAAAAGTTAAAGCAGATGCCA	429
464	QY	TTTACC CGGGATA CGGCTTCTGCTGAAATGCCAGCTTGC CGCGCAGTGC GGGA	523
430	DB	TTTACC CGGGATA CGGCTTCTGCTGAAATGCCAGCTTGC CGCGCAGTGC GGGA	489
524	QY	ACGGCATTTACTTTTATTGGCCCAACCCACAGAGTTCTTTGATCTCAACCGGTGATAGTCTC	583
490	DB	ACGGCATTTACTTTTATTGGCCCAACCCACAGAGTTCTTTGATCTCAACCGGTGATAGTCTC	549
584	QY	CGCGGTTAACCGCGCAGAGAGGCTGCTGCCAGTTTTCGCGGAATTCACCCCGACGA	643
550	DB	CGCGGTTAACCGCGCAGAGAGGCTGCTGCCAGTTTTCGCGGAATTCACCCCGACGA	609
644	QY	AAAACATCGATGAGATCGTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGG	703
610	DB	AAAACATCGATGAGATCGTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGG	669

[illegible]

Db	2830	ACCTCCTTGGTGGGTGTGGATTCAGCAGAGACTTTTGCTGCCGATCCACAAAGATGACGACA	2888
Qy	2924	TCCAGAGACTCTGTATCGCGGTTCCTTCGCGGGGAGCTTTGGTAACCTCCAGTGGCTGGC	2983
Db	2890	TCCAGACTCTGTATCGCGGTTCCTTCGCGGGGAGCTTTGGTAACCTCCAGTGGCTGGC	2949
Qy	2984	CAGAGCCACTGGCACCCGCGCACTGGAAAGGGCGCTCCGAAGGCAGGCACTCTGACGG	3043
Db	2950	CAGAGCCACTGGCGCACCCGCGCACTGGAAAGGGCGCTCCGAAGGCAGGCACTCTGACGG	3009
Qy	3044	AAGTTCTGAGGAAGCAGCGGGCACTCGAGCGCTGATGATTCGAAGGAACGTCGCAATA	3103
Db	3010	AAGTTCTGAGGAAGCAGCGGGCACTCGAGCGCTGATGATTCGAAGGAACGTCGCAATA	3069
Qy	3104	GCCTCAACCGCCTGCTGTCCCGAAGCCAAACGAGAGATTCTTCGAGCACCGTCGCGCT	3163
Db	3070	GCCTCAACCGCCTGCTGTTCGGAAGCCAAACGAGAGATTCTTCGAGCACCGTCGCGCT	3129
Qy	3164	TCGGCAACACTCTGGCTGGATGATCGTGAATTCCTTACGGCCTGGTCGAAGGCCGG	3223
Db	3130	TCGGCAACACTCTGGCTGGATGATCGTGAATTCCTTACGGCCTGGTCGAAGGCCGG	3189
Qy	3224	AGACTTGATTCGCGCTGCCAGATGTGGGCAACCCCACTGCTTGTTCGCTTGGATGCGATCT	3283
Db	3190	AGACTTGATTCGCGCTGCCAGATGTGGGCAACCCCACTGCTTGTTCGCTTGGATGCGATCT	3249
Qy	3284	CTGAGCCAGACGATAGGGTATGCGCAATGTTGTGGCCAAACGTCGAGCCAGATCCGCC	3343
Db	3250	CTGAGCCAGACGATAGGGTATGCGCAATGTTGTGGCCAAACGTCGAGCCAGATCCGCC	3309
Qy	3344	CAATGCGTGTGCGTGACCCGCTCCGTTGAGTCTGTCAACCGCAACCGCAGAAAGGCGAGTT	3403
Db	3310	CAATGCGTGTGCGTGACCCGCTCCGTTGAGTCTGTCAACCGCAACCGCAGAAAGGCGAGTT	3369
Qy	3404	CCTCCACAGGGGCCATGTTGTCGCACCATTCGCTGTGTGTTCACCGTCACTGTTGCTG	3463
Db	3370	CCTCCACAGGGGCCATGTTGTCGCACCATTCGCTGTGTGTTCACCGTCACTGTTGCTG	3429
Qy	3464	AAGGTGATGAGGTCAAGGCTTGGAGATGCAATCGCAATCATCGAGGCTATGAAGATGGAAG	3523
Db	3430	AAGGTGATGAGGTCAAGGCTTGGAGATGCAATCGCAATCATCGAGGCTATGAAGATGGAAG	3489
Qy	3524	CAACAAATCACTGCTTCTGTTGAGGGCAAAATCCGATCGGTTGTGGTTCCTGCTGCAACGA	3583
Db	3490	CAACAAATCACTGCTTCTGTTGAGGGCAAAATCCGATCGGTTGTGGTTCCTGCTGCAACGA	3549
Qy	3584	AGGTGGAAGTGGCGACTTGATCGTCTGTTTCCTAA	3621
Db	3550	AGGTGGAAGTGGCGACTTGATCGTCTGTTTCCTAA	3587

## RESULT, T 7

A97276	A97276	3728 bp	DNA	linear	PAT 26-JAN-2000
LOCUS					
DEFINITION	Sequence 1 from Patent WO99:18228.				

```

ACCESSION      A97276
VERSION        A97276.1  GI:6780661
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 3728)
Peters-Wendisch, P. and Eikmanns, B.
AUTHORS        METHOD FOR MICROBIAL PRODUCTION OF AMINO ACIDS OF THE ASPARTATE
TITLE
AND/OR GUANAMATE FAMILY AND AGENTS WHICH CAN BE USED IN SAID METHOD
JOURNAL        Patent: WO 9918228-A 1 15-APR-1999;
PETERS WENDISCH PETRA (DE); EIKMANN'S BERND (DE)
FEATURES
source
location/Qualifiers
1..3728
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

```

Query Match				Score 3578; DB 6; Length 3728;			
Best Local Similarity 100.0%; Pred. No. 0;							
Matches 3578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
QY	44	GCTTGAAGTCGTGACAGGTCAAGGAGTGTTCGCCGAAACAAATTCAGAGGAAACAAAC	103				
DB	10	GCTTGAAGTCGTGACAGGTCAAGGAGTGTTCGCCGAAACAAATTCAGAGGAAACAAAC	69				
QY	104	CGATGTTTGAATTCGGGGAATCGGGGTTTACGATCTAGGCGCAGTCACTGCTATCACCC	163				
DB	70	CGATGTTTGAATTCGGGGAATTCGGGTTTACGATCTAGGCGCAGTCACTGCTATCACCC	129				
QY	164	TTGCGGCTCTCTTGTGTAAGAAATTAATTAATCTAGTGTGCACTCAACATCTTTCAACGC	223				
DB	130	TTGCGGCTCTCTTGTGTAAGAAATTAATTAATCTAGTGTGCACTCAACATCTTTCAACGC	189				
QY	224	TTCCAGCATCAAAAGATCTTGTAGCAACCGCGGGAATCGCGTCCGTCTTTC	283				
DB	190	TTCCAGCATCAAAAGATCTTGTAGCAACCGCGGGAATTCGGTCCGTCTTTC	249				
QY	284	GTGACGACTCGAAACCGGTGCAGCCAGGTAGCTATTATCCCGGTGAAGATCGGGAT	343				
DB	250	GTGACGACTCGAAACCGGTGCAGCCAGGTAGCTATTATCCCGGTGAAGATCGGGAT	309				
QY	344	CAATTCACCGCTCTTTGCTTCTGAAGCTGTCCGATTTGTCGAGGCTCACCAGTCA	403				
DB	310	CAATTCACCGCTCTTTGCTTCTGAAGCTGTCCGATTTGTCGAGGCTCACCAGTCA	369				
QY	404	AGCGCTACTCGACATCGAATAATATCGGTGACGCTAAAGAAAGTTAAAGCAGATGCCA	463				
DB	370	AGCGCTACTCGACATCGAATAATATCGGTGACGCTAAAGAAAGTTAAAGCAGATGCCA	429				
QY	464	TTTACCGGGATACGGCTTCTGTCTGAAGTGTCCGATTTGTCGAGGCTCACCAGTCA	523				
DB	430	TTTACCGGGATACGGCTTCTGTCTGAAGTGTCCGATTTGTCGAGGCTCACCAGTCA	489				
QY	524	ACGGCATTAATTTATTTGCCCCAATTCGAGGTTCTTGATCTCAACCGGTGATAGTCTC	583				
DB	490	ACGGCATTAATTTATTTGCCCCAATTCGAGGTTCTTGATCTCAACCGGTGATAGTCTC	549				
QY	584	GGCGGTAAACCGCGGGAAGAGGTGTGTCCAGTTTGGCGAATCCACCCGAGCA	643				
DB	550	GGCGGTAAACCGCGGGAAGAGGTGTGTCCAGTTTGGCGAATCCACCCGAGCA	609				
QY	644	AAAAATCATGATGATCGTTAAAGCGCTGAAGGCGAGCTTACCCCATCTTTGTGAAGG	703				
DB	610	AAAAATCATGATGATCGTTAAAGCGCTGAAGGCGAGCTTACCCCATCTTTGTGAAGG	669				
QY	704	CAGTTGCCGCTGTGCGGACCGGATATGCGTTTGTGCTTCACTGATGAGCTTCGCA	763				
DB	670	CAGTTGCCGCTGTGCGGACCGGATATGCGTTTGTGCTTCACTGATGAGCTTCGCA	729				
QY	764	AATTAGCAACAGAGCATCTCTGAAGCTGAAGCGCTTTCGCGATGGCGGTATATG	823				
DB	730	AATTAGCAACAGAGCATCTCTGAAGCTGAAGCGCTTTCGCGATGGCGGTATATG	789				
QY	824	TGAAAGTCTGTGATTAACCTTCAGCATATTTGAAGTGCAGATCTCTGCGCATCACTG	883				
DB	790	TGAAAGTCTGTGATTAACCTTCAGCATATTTGAAGTGCAGATCTCTGCGCATCACTG	849				
QY	884	GAGAGTCTGACCTTTATGACGTGACTGCTCACTGACGCGTCTGACCAAAAGTTG	943				
DB	850	GAGAGTCTGACCTTTATGACGTGACTGCTCACTGACGCGTCTGACCAAAAGTTG	909				
QY	944	TCGAAATTCGCGCAGCAGCAATTTGGATCCAGAACTGCGTATCGCATTTGTGCGGATG	1003				
DB	910	TCGAAATTCGCGCAGCAGCAATTTGGATCCAGAACTGCGTATCGCATTTGTGCGGATG	969				
QY	1004	CAGTAAAGTCTGCCCTCCATTTGTTACGAGGCGCGGAAACCGTGGATTTCTTGTGTCG	1063				
DB	970	CAGTAAAGTCTGCCCTCCATTTGTTTACGAGGCGCGGAAACCGTGGATTTCTTGTGTCG	1029				

QY	1064	ATGAAAGGCGCAACCAACGCTCTTCTATCGAAATGAACCCACGTTATCCAGGTTGAGCAACCG	1123
DB	1030	ATGAAAGGCGCAACCAACGCTCTTCTATCGAAATGAACCCACGTTATCCAGGTTGAGCAACCG	1089
QY	1124	TGACTGAAAGAGTCAACGAGGTGGACCTTGTGAAGCGCAGATGCGTTGGTCTGTGTG	1183
DB	1090	TGACTGAAAGAGTCAACGAGGTGGACCTTGTGAAGCGCAGATGCGTTGGTCTGTGTG	1149
QY	1184	CAACCTTGAAGAAATTTGGTCTTGACCCCAAGATAAGATCAAGACCCACCGTGCAGCACTGC	1243
DB	1150	CAACCTTGAAGAAATTTGGTCTTGACCCCAAGATAAGATCAAGACCCACCGTGCAGCACTGC	1209
QY	1244	AGTGGCGCATCAACGAGGAGATCCAAACAAGCGCTTCGCCACAGATACCGGAATATCA	1303
DB	1210	AGTGGCGCATCAACGAGGAGATCCAAACAAGCGCTTCGCCACAGATACCGGAATATCA	1269
QY	1304	CGCGGTACCGCTCACACGAGCGGAGTGGCGTTTGTCTTGAACGCTCAAGCTCAAGCTCGGTG	1363
DB	1270	CGCGGTACCGCTCACACGAGCGGAGTGGCGTTTGTCTTGAACGCTCAAGCTCAAGCTCGGTG	1329
QY	1364	GCGAATCAACGCACTTTGATCTCATGCTGGTGAATGACCTGCCGTGGTTCGGACT	1423
DB	1330	GCGAATCAACGCACTTTGATCTCATGCTGGTGAATGACCTGCCGTGGTTCGGACT	1389
QY	1424	TTGAAACTCTGTGTCTGTCGTCACAGCGCGCTTGGCTGAGTTCAACCGTGTCTGSGTTG	1483
DB	1390	TTGAAACTCTGTGTCTGTCGTCACAGCGCGCTTGGCTGAGTTCAACCGTGTCTGSGTTG	1449
QY	1484	CAACCAACATTTGGTTTCTTGGTGGTGTCTGCGGAGAGGACTTCACTTCCAGCGCA	1543
DB	1450	CAACCAACATTTGGTTTCTTGGTGGTGTCTGCGGAGAGGACTTCACTTCCAGCGCA	1509
QY	1544	TCGCCACCGGATTCATTTGCCGATCACCCGACCTCTTTCAGGCTCCACCTGCTGATGATG	1603
DB	1510	TCGCCACCGGATTCATTTGCCGATCACCCGACCTCTTTCAGGCTCCACCTGCTGATGATG	1569
QY	1604	AGCAGGGAACGATCTCTGATTTACTTGGCAGATGTCAACCGTGAACAGCTCATGCTGTC	1663
DB	1570	AGCAGGGAACGATCTCTGATTTACTTGGCAGATGTCAACCGTGAACAGCTCATGCTGTC	1629
QY	1664	GTCCAAAGGATGTTGAGCTCTTATGATAAGCTGCTAAACATCAAGGATCTGCCACTGC	1723
DB	1630	GTCCAAAGGATGTTGAGCTCTTATGATAAGCTGCTAAACATCAAGGATCTGCCACTGC	1689
QY	1724	CACCGGTTTCCGTGACCGCTGAAAGAGTGTGGCCAGCGGCTTGTCTGCTGATCTCC	1783
DB	1690	CACCGGTTTCCGTGACCGCTGAAAGAGTGTGGCCAGCGGCTTGTCTGCTGATCTCC	1749
QY	1784	GTGAGCAGGACGACCTGGCAGTTTACTGATACACCTTCCGCGATGACACCACTCTTTC	1843
DB	1750	GTGAGCAGGACGACCTGGCAGTTTACTGATACACCTTCCGCGATGACACCACTCTTTC	1809
QY	1844	TTGCGACCCGAGTCCGCTCATTCGCACTGAAGCTTCGCGCAGAGCGGCTCGCAAACTGA	1903
DB	1810	TTGCGACCCGAGTCCGCTCATTCGCACTGAAGCTTCGCGCAGAGCGGCTCGCAAACTGA	1869
QY	1904	CTCTGAGCTTTTGTCCGTGGAGGCTTGGCGGCGGACCTTACGATGTGGCGATCGGTT	1963
DB	1870	CTCTGAGCTTTTGTCCGTGGAGGCTTGGCGGCGGACCTTACGATGTGGCGATCGGTT	1929
QY	1964	TCCTCTTTGAGGATCCGTTGGGACAGGCTCGACGAGTGTGCGCGAGGCGATGCCGAATGTA	2023
DB	1930	TCCTCTTTGAGGATCCGTTGGGACAGGCTCGACGAGTGTGCGCGAGGCGATGCCGAATGTA	1989
QY	2024	ACATTGAGTGTCTTCCGCGCGGCAACACCGTGGGATACACCCGTTACCCAGACTCCG	2083
DB	1990	ACATTGAGTGTCTTCCGCGCGGCAACACCGTGGGATACACCCGTTACCCAGACTCCG	2049
QY	2084	TCCTGCCGCGGTTTGTAAAGAAAGCTGCCAGCTCCGCGGTGGACATCTTCCGATCTTCG	2143
DB	2050	TCCTGCCGCGGTTTGTAAAGAAAGCTGCCAGCTCCGCGGTGGACATCTTCCGATCTTCG	2109
QY	2144	ACGCGCTTAAAGAGCTCTCCAGATGCGTCCAGCATCGACGAGCTCCTCGGAGCAACA	2203

Db 2110 ACAGCGCTTAAAGACGCTCTCCAGATGCGTCCAGCAATCGACGAGTCTCGAGACCAACA 2169  
Qy 2204 CCGCGGTAGCGAGGTGGCTATGCTGCTATCTCTGATCTCTGATCCAAATGAAGAAGC 2263  
Db 2170 CCGCGGTAGCGAGGTGGCTATGCTGCTATCTCTGATCTCTGATCCAAATGAAGAAGC 2229  
Qy 2264 TCTACACCTCTGATTAATCTACCTAAAGATGCGAGAGAGATCGTCAAGTCTGCGCTCA 2323  
Db 2230 TCTACACCTCTGATTAATCTACCTAAAGATGCGAGAGAGATCGTCAAGTCTGCGCTCA 2289  
Qy 2324 TCTTGGCCATTAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2383  
Db 2290 TCTTGGCCATTAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2349  
Qy 2384 CCGCACTGCGCGGTGAATTCGATCTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 2443  
Db 2350 CCGCACTGCGCGGTGAATTCGATCTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 2409  
Qy 2444 GCGAGCTGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2503  
Db 2410 GCGAGCTGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2469  
Qy 2504 CCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2563  
Db 2470 CCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2529  
Qy 2564 GCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2623  
Db 2530 GCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2589  
Qy 2624 GCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2683  
Db 2590 GCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2649  
Qy 2684 GCGCTTACCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2743  
Db 2650 GCGCTTACCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2709  
Qy 2744 CACTGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2803  
Db 2710 CACTGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2769  
Qy 2804 TGCTGGGAGCGCCAAACCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2863  
Db 2770 TGCTGGGAGCGCCAAACCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2829  
Qy 2864 ACCTCGTGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2923  
Db 2830 ACCTCGTGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2889  
Qy 2924 TCCGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2983  
Db 2890 TCCGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2949  
Qy 2984 CAGAGCCACTGCGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3043  
Db 2950 CAGAGCCACTGCGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3009  
Qy 3044 AAGTTCTCTGAGGAG 3103  
Db 3010 AAGTTCTCTGAGGAG 3069  
Qy 3104 GCCTCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3163  
Db 3070 GCCTCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3129  
Qy 3164 TCGGCAACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3223  
Db 3130 TCGGCAACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3189  
Qy 3224 AGACTTTGATCGGCTGCGAGATGTCGCGACCGCACTGCTGCTGCTGCTGCTGCTGCTGCT 3283

Db 3190 AGACTTTGATCGGCTGCGAGATGTCGCGACCGCACTGCTGCTGCTGCTGCTGCTGCTGCT 3249  
Qy 3284 CTGAGCCAGACGATTAAGGGTATGCGCAATGTTGTGGCAACGTCACCGCCAGATCCGCC 3343  
Db 3250 CTGAGCCAGACGATTAAGGGTATGCGCAATGTTGTGGCAACGTCACCGCCAGATCCGCC 3309  
Qy 3344 CAATGCTGTCGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3403  
Db 3310 CAATGCTGTCGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3369  
Qy 3404 CCTCCAAACAAGGGCCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3463  
Db 3370 CCTCCAAACAAGGGCCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3429  
Qy 3464 AAGTGATGAGTCAAGGGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3523  
Db 3430 AAGTGATGAGTCAAGGGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3489  
Qy 3524 CAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCTGCTGCTGCTGCTGCTGCTGCT 3583  
Db 3490 CAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCTGCTGCTGCTGCTGCTGCTGCT 3549  
Qy 3584 AGTGGAAGTGGCGACTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3621  
Db 3550 AGTGGAAGTGGCGACTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3587

RESULT 8  
BD140686 3728 bp DNA linear PAT 18-SEP-2002  
LOCUS Method for microbial production of amino acids of the aspartate  
DEFINITION and/or glutamate family and agents which can be used in said  
method.  
ACCESSION BD140686  
VERSION BD140686.1 GI:23235631  
KEYWORDS JP 2002508921-A/1.  
SOURCE Corynebacterium glutamicum  
ORGANISM Corynebacterium glutamicum  
REFERENCE Bixmanns.B., Wendisch,P.P. and Sahm.H.  
AUTHORS Method for microbial production of amino acids of the aspartate  
TITLE and/or glutamate family and agents which can be used in said method  
JOURNAL Patent: JP 2002508921-A 1 26-MAR-2002;  
COMMENT FORSCHUNGSZENTRUM JULICH GMBH  
OS Corynebacterium glutamicum  
PN JP 2002508921-A/1  
PD 26-MAR-2002  
PF 30-SEP-1998 JP 2000515022  
PR 04-OCT-1997 DE 197 43 894.6.14-JUL-1998 DE 198 31 609.7 PI  
BERND BIXMANN, PETRA PETERS WENDISCH, HERMANN SAHM PC  
C12N15/09,C12N1/21,C12F13/20// (C12P13/20,C12R1:15),C12N15/00 CC  
pyruvate carboxylase gene  
FH Key Location/Qualifiers  
FT CDS (165).. (3587).

FEATURES  
1. 3728  
/organism="Corynebacterium glutamicum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1718"

ORIGIN  
Query Match 98.8%; Score 3578; DB 6; Length 3728;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 GCTTGAAGTCTGTCAGGTGTCAGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAAAAC 103  
Db 10 GCTTGAAGTCTGTCAGGTGTCAGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAAAAC 69  
Qy 104 CCATGTTTATTGGGGGAATCGGGGTTACCATATAGGACGAGTGTGCTATCACCC 163

Db 70 CGATGTTTGAATGGGGAAATCGGGGTTACGATACCTAGGACGCACTGACTGCTATCACCC 129  
QY 164 TTGGGGTCTCTTTGTTGAAAGAAATAATTACTCTAGTGTGACATCACAATCTTCAACGC 223  
Db 130 TTGGGGTCTCTTTGTTGAAAGAAATAATTACTCTAGTGTGACATCACAATCTTCAACGC 189  
QY 224 TTCCAGCATTTCAAAAGATTTTGTGTAGCAAAACCGCGGGAATAATCGCGTCCGTCTTCC 283  
Db 190 TTCCAGCATTTCAAAAGATTTTGTGTAGCAAAACCGCGGGAATAATCGCGTCCGTCTTCC 249  
QY 284 GTGCGACATCGAATACCGGTGACGACCACTGATGATTTATCCCGTCAAGATCGGGAT 343  
Db 250 GTGCGACATCGAATACCGGTGACGACCACTGATGATTTATCCCGTCAAGATCGGGAT 309  
QY 344 CATTTCCACGCTCTTTTGTCTTCTGAAGCTGTCCGCAITTTGTACCGAAGGCTCACAGTCA 403  
Db 310 CATTTCCACGCTCTTTTGTCTTCTGAAGCTGTCCGCAITTTGTACCGAAGGCTCACAGTCA 369  
QY 404 AGCGTACCTGACATCGATGAATTTATCGGTGACGCTTAAAGTTAAAGCAGATGCCA 463  
Db 370 AGCGTACCTGACATCGATGAATTTATCGGTGACGCTTAAAGTTAAAGCAGATGCCA 429  
QY 464 TTATCCCGGATACGCTTCTGTCTGAAATGCCCCAGCTTGCCTGCGAGTGTGCGAAA 523  
Db 430 TTATCCCGGATACGCTTCTGTCTGAAATGCCCCAGCTTGCCTGCGAGTGTGCGAAA 489  
QY 524 ACGGCAATTAATTTATGCCCCAAACCCGAGAGTTCTGATCTCACCGGTGATAGTCTC 583  
Db 490 ACGGCAATTAATTTATGCCCCAAACCCGAGAGTTCTGATCTCACCGGTGATAGTCTC 549  
QY 584 GCGCGTAACCGCGGAAAGAGTGTGTCTCCAGTTTGTGGCGAATCCACCCCGAGCA 643  
Db 550 GCGCGTAACCGCGGAAAGAGTGTGTCTCCAGTTTGTGGCGAATCCACCCCGAGCA 609  
QY 644 AAAACATCGATGATCGTTAAAGCGCTGAAGGCGAGCTTACCCCATCTTTGTGAAGG 703  
Db 610 AAAACATCGATGATCGTTAAAGCGCTGAAGGCGAGCTTACCCCATCTTTGTGAAGG 669  
QY 704 CAGTTCCCGTGTGCGAGCGAGTATGCGTTTGTGTCTTCACTGATGAGCTTCGCA 763  
Db 670 CAGTTCCCGTGTGCGAGCGAGTATGCGTTTGTGTCTTCACTGATGAGCTTCGCA 729  
QY 764 AATTAGCAACAGAGCATCTGTGAAGCTGAAGCGGCTTTCGGGATGCGCGGTATATG 823  
Db 730 AATTAGCAACAGAGCATCTGTGAAGCTGAAGCGGCTTTCGGGATGCGCGGTATATG 789  
QY 824 TCGAAGTCTGTGATTAACCTCAGCATATGAAGTGCAGATCTTGCAGATCACACTG 883  
Db 790 TCGAAGTCTGTGATTAACCTCAGCATATGAAGTGCAGATCTTGCAGATCACACTG 849  
QY 884 GAGAACTTGTACCTTTATGAACGTGACTGCTCACTGACGCTGCTCAACCAAAAGTTG 943  
Db 850 GAGAACTTGTACCTTTATGAACGTGACTGCTCACTGACGCTGCTCAACCAAAAGTTG 909  
QY 944 TCGAAATTCGCGCAGCAGACATTTGGATTCAGACTGCGTATGCAATTTGTGGGATG 1003  
Db 910 TCGAAATTCGCGCAGCAGACATTTGGATTCAGACTGCGTATGCAATTTGTGGGATG 969  
QY 1004 CAGTAAAGTTCTCCGCTCCATTTGTTTACCGGCGCGGAAACCGTGAATTTCTTGGTCG 1063  
Db 970 CAGTAAAGTTCTCCGCTCCATTTGTTTACCGGCGCGGAAACCGTGAATTTCTTGGTCG 1029  
QY 1064 ATGAAAGGCGCACACGCTTTCATCGAATGAAACCGAGTATCCAGTTGAGCACCG 1123  
Db 1030 ATGAAAGGCGCACACGCTTTCATCGAATGAAACCGAGTATCCAGTTGAGCACCG 1089  
QY 1124 TGAATGAAGAGTCAACCGAGGTGACCTGTGTAAGGCGCAGATGCGCTTGGCTGTGTTG 1183  
Db 1090 TGAATGAAGAGTCAACCGAGGTGACCTGTGTAAGGCGCAGATGCGCTTGGCTGTGTTG 1149  
QY 1184 CAACCTTGAAGGAATTTGGTCTGACCCCAAGATGAAGATCAAGACCCACGTCAGACATGC 1243  
Db 1150 CAACCTTGAAGGAATTTGGTCTGACCCCAAGATGAAGATCAAGACCCACGTCAGACATGC 1209

QY 1244 AGTGCCGATCACCCAGGAAGATCCAAAACAAGGCTTCCGCCAGATACCGGAATATCA 1303  
Db 1210 AGTGCCGATCACCCAGGAAGATCCAAAACAAGGCTTCCGCCAGATACCGGAATATCA 1269  
QY 1304 CCGGTAACGCTCACCAAGCGGAGTGGGTTTGTCTTGAAGTGCAGTGCAGTCCGTTG 1363  
Db 1270 CCGGTAACGCTCACCAAGCGGAGTGGGTTTGTCTTGAAGTGCAGTGCAGTCCGTTG 1329  
QY 1364 GCGAAATCACCGCACACTTTTGTCTTGAATGAAATGAAATGAAATGAAATGAAATG 1423  
Db 1330 GCGAAATCACCGCACACTTTTGTCTTGAATGAAATGAAATGAAATGAAATGAAATG 1389  
QY 1424 TTGAAATCTGCTG 1483  
Db 1390 TTGAAATCTGCTG 1449  
QY 1484 CAACCAACATTTGTTTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1543  
Db 1450 CAACCAACATTTGTTTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1509  
QY 1544 TCGCCACCGGATTCATTTGCGGATCACCGGACCTCTTCAAGGCTCACTTCCAAAGCGCA 1603  
Db 1510 TCGCCACCGGATTCATTTGCGGATCACCGGACCTCTTCAAGGCTCACTTCCAAAGCGCA 1569  
QY 1604 AGCAGGAGCATCTTGGATTTACTTGGCAGATGTCAACCGTGAACAGCTTCATGTGTGC 1663  
Db 1570 AGCAGGAGCATCTTGGATTTACTTGGCAGATGTCAACCGTGAACAGCTTCATGTGTGC 1629  
QY 1664 GTCCAAAGGATTTGAGCTCTTATCGATGAAGTGTCTTAAATCAAGGATCTGCCACTGC 1723  
Db 1630 GTCCAAAGGATTTGAGCTCTTATCGATGAAGTGTCTTAAATCAAGGATCTGCCACTGC 1689  
QY 1724 CAGCGGTTCCCGTGAACGCTTGAAGCAGCTTGGCCAGCGGCTTGGTGTGTGTGTGTGTGTG 1783  
Db 1690 CAGCGGTTCCCGTGAACGCTTGAAGCAGCTTGGCCAGCGGCTTGGTGTGTGTGTGTGTG 1749  
QY 1784 GTGAGCAGCAGCATGCGGATTTACTGATACCACTTCCCGGATGACACCACTCTTTCG 1843  
Db 1750 GTGAGCAGCAGCATGCGGATTTACTGATACCACTTCCCGGATGACACCACTCTTTCG 1809  
QY 1844 TTGCGACCGGATTCGCTCATTTGCGCATGAGCTTGGCCAGCGGCTTGGCAGGCTGCA 1903  
Db 1810 TTGCGACCGGATTCGCTCATTTGCGCATGAGCTTGGCCAGCGGCTTGGCAGGCTGCA 1869  
QY 1904 CTCTGAGCTTTTGTCCGTGGAGGCTTGGCGGCGGCGGCTTACGATGTGGCGATCGCTT 1963  
Db 1870 CTCTGAGCTTTTGTCCGTGGAGGCTTGGCGGCGGCGGCTTACGATGTGGCGATCGCTT 1929  
QY 1964 TCCTCTTTGAGGATTCGCTGGGACAGGCTGACGCTGCGGAGGCGGATGCGGAATGTAA 2023  
Db 1930 TCCTCTTTGAGGATTCGCTGGGACAGGCTGACGCTGCGGAGGCGGATGCGGAATGTAA 1989  
QY 2024 ACATTCAGATGCTGCTTCCGCGCGGCAACACCGTGGGATACACCCGTACCCAGATCCG 2083  
Db 1990 ACATTCAGATGCTGCTTCCGCGCGGCAACACCGTGGGATACACCCGTACCCAGATCCG 2049  
QY 2084 TCTGCGCGGCTTGTGTTAAGGAAGCTCCAGCTCCGCGGTGAGCATCTTCCGATCTTTCG 2143  
Db 2050 TCTGCGCGGCTTGTGTTAAGGAAGCTCCAGCTCCGCGGTGAGCATCTTCCGATCTTTCG 2109  
QY 2144 ACGGCTTAAACGAGCTCTCCAGATGCTCCAGCAATTCAGCAGCTTCCGAGACCAACA 2203  
Db 2110 ACGGCTTAAACGAGCTCTCCAGATGCTCCAGCAATTCAGCAGCTTCCGAGACCAACA 2169  
QY 2204 CCGCGGTAGCGAGGTGGCTATGCTTATTTCTGCTGATCTCTGATTCGGAATGAAAGC 2263  
Db 2170 CCGCGGTAGCGAGGTGGCTATGCTTATTTCTGCTGATCTCTGATTCGGAATGAAAGC 2229  
QY 2264 TCTACACCTTGAATTAACCTTAAGATGACAGGAGATCGCTCAAGTCTGCGGCTCACA 2323  
Db 2230 TCTACACCTTGAATTAACCTTAAGATGACAGGAGATCGCTCAAGTCTGCGGCTCACA 2289



2324 TCCTGGCCATTAAAGGATATGGCTGGTCTGCTTCCGCCAGCTGCGGTAAACCAAGCTGGTCA 2383  
2290 TCCTGGCCATTAAAGGATATGGCTGGTCTGCTTCCGCCAGCTGCGGTAAACCAAGCTGGTCA 2349  
2384 CGGACATGCGCGCGTGAATTCGATCTGCGAGTGCAGCTGCACACCCACGACACTGCGGGTG 2443  
2350 CGGACATGCGCGCGTGAATTCGATCTGCGAGTGCAGCTGCACACCCACGACACTGCGGGTG 2409  
2444 GCCAGCTGGCAACCTACTTTGCTGCGAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGT 2503  
2410 GCCAGCTGGCAACCTACTTTGCTGCGAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGT 2469  
2504 CGGACATGCTGCGGACACTTCCAGCCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2563  
2470 CGGACACCTGCTGCGGACACTTCCAGCCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 2529  
2564 CGCACACCGCTGCGGATACCGGTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGCTACT 2623  
2530 CGCACACCGCTGCGGATACCGGTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGCTACT 2589  
2624 GGGAGCAGTGGCGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2683  
2590 GGGAGCAGTGGCGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2649  
2684 GCGTCTACCGCCAGCAATCCAGCGGACAGTGTCCAACTCGCTGCTGCTGCTGCTGCTGCTG 2743  
2650 GCGTCTACCGCCAGCAATCCAGCGGACAGTGTCCAACTCGCTGCTGCTGCTGCTGCTGCTG 2709  
2744 CACTGGGCGCTTGGGATGCTTGAATCTCATCGAGACAACTAGCGAGCGGTTAATGAGA 2803  
2710 CACTGGGCGCTTGGGATGCTTGAATCTCATCGAGACAACTAGCGAGCGGTTAATGAGA 2769  
2804 TGCTGGGACGCGCAACCAAGTCAACCTCTCCAAAGTGTGTCGAGCTCGCACTCGCACTCC 2863  
2770 TGCTGGGACGCGCAACCAAGTCAACCTCTCCAAAGTGTGTCGAGCTCGCACTCGCACTCC 2829  
2864 ACCTGTTGTTGCGGGTGTGATCCAGCAGACTTGTGTCGCGATTCACAAAAGTACGACA 2923  
2830 ACCTGTTGTTGCGGGTGTGATCCAGCAGACTTGTGTCGCGATTCACAAAAGTACGACA 2889  
2924 TCCAGACTCTGCTATCGCTTCTGCGCGGAGCTTGTGTAACCTCTCCAGGTTGGCTGGC 2983  
2890 TCCAGACTCTGCTATCGCTTCTGCGCGGAGCTTGTGTAACCTCTCCAGGTTGGCTGGC 2949  
2984 CAGAGCCACTCGCACCCCGGCACTGGAAGCGCGCTCGAAGCAAGGCAAGGCAAGGCAAGG 3043  
2950 CAGAGCCACTCGCACCCCGGCACTGGAAGCGCGCTCGAAGCAAGGCAAGGCAAGGCAAGG 3009  
3044 AAGTCTCTGAGGAGAGCAGCGCACCTCGAGCTGATGATTCACAGGACGCTGCGCAATA 3103  
3010 AAGTCTCTGAGGAGAGCAGCGCACCTCGAGCTGATGATTCACAGGACGCTGCGCAATA 3069  
3104 GCCTCAACCGCGCTGCTGTTCCGGAAGCAACCGAAGAGTTCCTCGAGCAGCGCTCGCGCT 3163  
3070 GCCTCAACCGCGCTGCTGTTCCGGAAGCAACCGAAGAGTTCCTCGAGCAGCGCTCGCGCT 3129  
3164 TCGGCAACACTCTGCGCTGATGATCTGTAATCTTCTAGCGCTGCTGTAAGCGCGCG 3223  
3130 TCGGCAACACTCTGCGCTGATGATCTGTAATCTTCTAGCGCTGCTGTAAGCGCGCG 3189  
3224 AGACTTTGATCGCGCTGCGATGTCGACCGCAACCGAAGAGTTCCTCGAGCAGCGCTCGCGCT 3283  
3190 AGACTTTGATCGCGCTGCGATGTCGACCGCAACCGAAGAGTTCCTCGAGCAGCGCTCGCGCT 3249  
3284 CTGAGCCAGACGATAGGGTATGCGCAATGTTGTCGCAAGCTGTAAGCGCGAGTTCGCGC 3343  
3250 CTGAGCCAGACGATAGGGTATGCGCAATGTTGTCGCAAGCTGTAAGCGCGAGTTCGCGC 3309  
3344 CAATGCGTGTGCGTGAACCGCTCGGTTGAGTGTGTGTCGCGCAACCGCAAGGAGGAGG 3403  
3310 CAATGCGTGTGCGTGAACCGCTCGGTTGAGTGTGTGTCGCGCAACCGCAAGGAGGAGG 3369  
3404 CCTCCAAACAGGGCCATGTTGCTGCAACCATTCGCTGGTGTGTGCTACCGCTGACTGTTGCTG 3463

Db 3370 CCTCCAAAGAGGCCCATGTTGCTGCGACCATTCCTGGTGTGTCACCGTGAATGTTGCTG 3429  
Qy 3464 AAGTGATGAGGCTCAAGGCTGGAGATGCGAGTGCAGTCAATCATCGAGGCTATGAAGATGAAG 3523  
Db 3430 AAGTGATGAGGCTCAAGGCTGGAGATGCGAGTGCAGTCAATCATCGAGGCTATGAAGATGAAG 3489  
Qy 3524 CAACAATCACTGCTTCTGTTGACGGCAAAATCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3583  
Db 3490 CAACAATCACTGCTTCTGTTGACGGCAAAATCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3549  
Qy 3584 AGTGGAGGCTGGGCACTTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3621  
Db 3550 AGTGGAGGCTGGGCACTTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3587  
  
RESULT 9  
AF503915 3657 bp DNA linear BCT 13-MAY-2002  
LOCUS Corynebacterium crenatum strain CD945 pyruvate carboxylase (pyc)  
DEFINITION gene, complete cds.  
ACCESSION AF503915  
VERSION AF503915.1 GI:20531756  
KEYWORDS  
SOURCE Corynebacterium crenatum  
ORGANISM Corynebacterium crenatum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
REFERENCE 1 (bases 1 to 3657)  
AUTHORS Wang, J., Ding, J. and Liu, Y.  
TITLE Cloning and Expression of Pyruvate Carboxylase Gene in Corynebacterium crenatum CD945  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3657)  
AUTHORS Wang, J., Ding, J. and Liu, Y.  
TITLE Direct Submission  
JOURNAL Submitted (18-APR-2002) Center of Microbial Biotechnology,  
Institute of Microbiology, Chinese Academy of Sciences, Beijing  
100080, China  
FEATURES  
source  
1. 3657  
/organism="Corynebacterium crenatum"  
/mol\_type="genomic DNA"  
/strain="CD945"  
/db\_xref="taxon:168810"  
151..3573  
/gene="pyc"  
151..3573  
/gene="pyc"  
/codon\_start=1  
/translation="pyruvate carboxylase"  
/protein\_id="AA027458.1"  
/db\_xref="GI:20531756"  
PDB:1K9A  
/translation="MSTHTSTLPFKILVANRGEIATVAFRAALETGAATVATYPR  
EDRGFHSFASAEVAVRIGTEGSPVKAYLIDIEIIIAKKKADAIYFGYGLFSENAQL  
AREPNTGIVGTPTEVLDITDKSRVATKAAGLFLAEPSTSKNIDIDIVKSAEG  
QTCYLFVAVAGGGRGMRVSSPDELKLALEARAEAAFGDGSVVERAVINPQH  
TEVQILGRTGEVHLVERDCLORRHKVVEIAPAOHLDPDLRDRICADAVKFCRSI  
GYQAGTEVFLVDEKGNVFIEMPRIOVEHTVEEVDLVKQWELAAAGATLKEI  
GUTQDKITKHGAALQCRITTEDPNPFRPTGTITATYRSPGAGVRLDGAQLGGEIT  
AHFDSMLVKMTCCGSDSFETAVARALAEFTVSGVATNIGFLRALLREEDFTSKRIA  
TGFI GDHPLHLQAPADDEQGRILYLDVTVNPHGVPRKDVAAPIIDKLNKIDPL  
PGSDRLKQLGPAARFADREQDALVDTTDFDAHQSLIATVRSFALPKPAEAVA  
KLTPILLSVEAGGATYDVMRFLEPDPMDRLDELREAMPVNIQMLRGRNTVGYTP  
VPSKCRAPVKEATSGVDIPRI FDALNDVSMRPAIDAVLETNTAVAEVAMASGDL  
SDPNKELTYLDYLLWASEIIVKSGAHILAKDMAGLLRPAATKLVATLRRERDLPVH  
VHTDTAGQLATYFAAQAQADAVDGASAPLSGTSQPSLSAIVAAFAHTRDGLS  
LEAVSDLPYWEAVGLYLPESGTPGPTGVIRHEIFGGQLSNLRQAATGLADRF  
ELIEDNYAAVEMLEGRPTKVTPSSKVGDLALHLVGAVDPADFAADPKQKIDPDSVI  
AFRLGELGNPPGWPPELRLTALGREGKAPLFEVPEEQHLDADDSKERRSLNR  
LLFPKPTFELEHRRFNGTISALDRREFYGLVEGRETILIRLPDVRTLLRLDAISE  
PDDKGMNVNANVGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFAGVVTVA



EGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVPAATKVEGDLIVVIS"

ORIGIN

Query Match 95.8%; Score 3469; DB 1; Length 3657;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 3508; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 49 AAGTCGTGAGTCAAGGAGTGTTCGCCGAAACATTTGAGAGAAACAAACCCGATG 108  
DB 1 AAGTCGTGAGTCAAGGAGTGTTCGCCGAAACATTTGAGAGAAACAAACCCGATG 60  
QY 109 TTTGATTTGGGGAAATCGGGGGTTACGATACAGGAGCGAGTCACTGCTATCACCTTTGGC 168  
DB 61 TTTGATTTGGGGAAATCGGGGGTTACGATACAGGAGCGAGTCACTGCTATCACCTTTGGT 120  
QY 169 GGTCTCTTTGTTGAAAGGAATTAATTAATCTAGTGTGCACTCAATCTTTCAACGCTTCCA 228  
DB 121 GGTCTCTTTGTTGAAAGGAATTAATTAATCTAGTGTGCACTCAATCTTTCAACGCTTCCA 180  
QY 229 GCATTCAAAAGATCTTGGTAGCAACCGCGCGAAATCGCGGTCCGCTTTCCGTTGCA 288  
DB 181 GCATTCAAAAGATCTTGGTAGCAACCGCGCGAAATCGCGGTCCGCTTTCCGTTGCA 240  
QY 289 GCATTCGAAACCGGTGACGCCACCGTAGCTATTTACCCCGTGAAGATTCGGGATCAATTC 348  
DB 241 GCATTCGAAACCGGTGACGCCACCGTAGCTATTTACCCCGTGAAGATTCGGGATCAATTC 300  
QY 349 CACCGCTCTTTTCTTCTGAGCTCTCGGCTGCTGAGTGTGAGGCTCAACGATCAAGCG 408  
DB 301 CACCGCTCTTTTCTTCTGAGCTCTCGGCTGCTGAGTGTGAGGCTCAACGATCAAGCG 360  
QY 409 TACCTGGACATCGATGAAATTAATCGGTGAGCTAAAGGTTAAAGCAGATGCCATTTAC 468  
DB 361 TACCTGGACATCGATGAAATTAATCGGTGAGCTAAAGGTTAAAGCAGATGCCATTTAC 420  
QY 469 CCGGATACGGTCTTCTGCTGAAATTCGCCAGCTGTCGAGTGTGCGGAGTGTGCGGAAACGGC 528  
DB 421 CCGGATACGGTCTTCTGCTGAAATTCGCCAGCTGTCGAGTGTGCGGAGTGTGCGGAAACGGC 480  
QY 529 ATTACTTTTATTTGGCCCAACCCAGAGTTCCTGATCTCACCGGTGATAAGTCTCGTGG 588  
DB 481 ATTACTTTTATTTGGCCCAACCCAGAGTTCCTGATCTCACCGGTGATAAGTCTCGTGG 540  
QY 589 GTAAACCGCGCGAAGAGCTGCTGCGAGTTTGGCGGAATTCACCCCGAGCAAAAC 648  
DB 541 GTAAACCGCGCGAAGAGCTGCTGCGAGTTTGGCGGAATTCACCCCGAGCAAAAC 600  
QY 649 ATCGATGACATCGTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAGCGAGTT 708  
DB 601 ATCGATGACATCGTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAGCGAGTT 660  
QY 709 GCGGTGTGTGGCGAGCGGATGCGTTTGTGCTTCACTGATGAGCTTCGCAATTA 768  
DB 661 GCGGTGTGTGGCGAGCGGATGCGTTTGTGCTTCACTGATGAGCTTCGCAAAATG 720  
QY 769 GCAACAGAGCATCTGTGAAGCTGAAGCGCTTTTCGCGATGCGCGGTATATGCGAA 828  
DB 721 GCAACAGAGCATCTGTGAAGCTGAAGCGCAATTCGCGAGCGTTTGTGATGTCGAA 780  
QY 829 CGTGTGTGATTAACCTCAGCATATGGAAGTGCAGATCTTTGGCGATCACACTGGAGAA 888  
DB 781 CGTGTGTGATTAACCCCGAGCAGATGGAAGTGCAGATCTTTGGCGATCGCACTGGAGAA 840  
QY 889 GTTGTACACTTATGAGCTGCTGCTACCTGCGCGTCTGCTACCAAAAGTTGTGAA 948  
DB 841 GTTGTACACTTATGAGCTGCTGCTACCTGCGAGCGTCTGCTACCAAAAGTTGTGAA 900  
QY 949 ATTGCGCCAGCAGCATTTGGATCCAGAACTGCGGTGATGCGATTTGTGCGGATGCACTA 1008  
DB 901 ATTGCGCCAGCAGCATTTGGATCCAGAACTGCGGTGATGCGATTTGTGCGGATGCACTA 960  
QY 1009 AAGTTCTGCGCTCCATTTGGTACAGGCGGGGAAACCGTGGATTTCTTGTTCGATGAA 1068

DB 961 AAGTTCTGCGCTCCATTTGGTTACAGGCGCGGGAAACCGTGGAAATTTCTTGTTCGATGAA 1020  
QY 1069 AAGGCAACACACGTCTTTCATCGAAATGAACCCATATCCAGTTTGAGCACACCGTGACT 1128  
DB 1021 AAGGCAACACACGTCTTTCATCGAAATGAACCCATATCCAGTTTGAGCACACCGTGACT 1080  
QY 1129 GAAGAGTCAACGAGTGGAGCTGTGAGGCGGAGATGCGCTTGTGCTGTGCTGCTGCTCAACC 1188  
DB 1081 GAAGAGTCAACGAGTGGAGCTGTGAGGCGGAGATGCGCTTGTGCTGTGCTGCTGCTCAACC 1140  
QY 1189 TTGAAGGAATTTGGTGTGACCCCAAGATGAAGATCAAGACCCAGGTGCGAGCTGCAAGTGC 1248  
DB 1141 TTGAAGGAATTTGGTGTGACCCCAAGATGAAGATCAAGACCCAGGTGCGAGCTGCAAGTGC 1200  
QY 1249 CGCATCAACAGGAGATCCAAACAAACCGCTTCGCGCCAGATACCGGACTATCAACCGG 1308  
DB 1201 CGCATCAACAGGAGATCCAAACAAACCGCTTCGCGCCAGATACCGGACTATCAACCGG 1260  
QY 1309 TACCGTCTACCAAGGCGAGCTGCGCTTCTGCTTGTGAGCGGTGCGAGCTCAGTCCGTTGCGGAA 1368  
DB 1261 TACCGTCTACCAAGGCGAGCTGCGCTTCTGCTTGTGAGCGGTGCGAGCTCAGTCCGTTGCGGAA 1320  
QY 1369 ATCACCGCACTTTGACTCTCATGCTGTGAAATGACCTGCGGTGCTTCCGACTTTGAA 1428  
DB 1321 ATCACCGCACTTTGACTCTCATGCTGTGAAATGACCTGCGGTGCTTCCGACTTTGAA 1380  
QY 1429 ACTGCTGTGCTGCTGCAAGCGCGCTTGGTGTGAGTTCAACGCTGTCTGCTGTGCAACC 1488  
DB 1381 ACTGCTGTGCTGCTGCAAGCGCGCTTGGTGTGAGTTCAACGCTGTCTGCTGTGCAACC 1440  
QY 1489 AACATTTGGTCTTGTGCTGCTGCGGAGAGGACTTCACTTCCAAAGCGCATTCGCC 1548  
DB 1441 AACATTTGGTCTTGTGCTGCTGCTGCGGAGAGGACTTCACTTCCAAAGCGCATTCGCC 1500  
QY 1549 ACCGGAATTCATTTGCCGATCACCGCGACCTTCCAGCTTCCAGCTGCTGATGAGCAG 1608  
DB 1501 ACCGGAATTCATTTGCCGATCACCGCGACCTTCCAGCTTCCAGCTGCTGATGAGCAG 1560  
QY 1609 GAGCGCATCTGATGATTTGAGCAGATGTCACCGTGAACCAAGCTCATGCTGCTGCTGCTCA 1668  
DB 1561 GAGCGCATCTGATGATTTGAGCAGATGTCACCGTGAACCAAGCTCATGCTGCTGCTGCTCA 1620  
QY 1669 AAGGATTTGCGAGCTCCTATCGATAGCTGCTTAACTCAAGGATGCTGCTGCTGCTGCTGCT 1728  
DB 1621 AAGGATTTGCGAGCACTTATGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
QY 1729 GGTTCCTGCTGAGCGCTGAGCAGCTTGGCCAGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCT 1788  
DB 1681 GGTTCCTGCTGAGCGCTGAGCAGCTTGGCCAGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCT 1740  
QY 1789 CAGGACGCTGCGAGCTTACTGATACCACTTCCGCGATGCAACCGAGTCTTGTGCTGCTGCT 1848  
DB 1741 CAGGACGCTGCGAGCTTACTGATACCACTTCCGCGATGCAACCGAGTCTTGTGCTGCTGCT 1800  
QY 1849 ACCCGAGTCCGCTCATTTCCGACTGAGGCTGCGGAGAGCGCTGCGAAAGCTGACTCTCT 1908  
DB 1801 ACCCGAGTCCGCTCATTTCCGACTGAGGCTGCGGAGAGAGCGCTGCGAAAGCTGACTCTCT 1860  
QY 1909 GAGCTTTGCTGAGGCTTGGGCGGCGGAGCACTTACGATGCTGCGGATGCTGCTTCTCTC 1968  
DB 1861 GAGCTTTGCTGAGGCTTGGGCGGCGGAGCACTTACGATGCTGCGGATGCTGCTTCTCTC 1920  
QY 1969 TTTGAGGATCCGTTGAGGCTGCGAGCTGCGGAGGCGATGCGGATGCTGCTGCTGCTGCT 2028  
DB 1921 TTTGAGGATCCGTTGAGGCTGCGAGCTGCGGAGGCGATGCGGATGCTGCTGCTGCTGCT 1980  
QY 2029 CAGATGCTGCTGCGGCGCGCAACACCGTGGGATACACCGGATACCGGACTCCGCTGCTG 2088  
DB 1981 CAGATGCTGCTGCGGCGCGCAACACCGTGGGATACACCGGATACCGGACTCCGCTGCTG 2040  
QY 2089 CCGCGCTTTGTTAAAGGAGCTGCGAGCTGCGGCGTGGACATCTTCCGCTCTTCCGAGCGG 2148  
DB 2041 CCGCGCTTTGTTAAAGGAGCTGCGAGCTGCGGCGTGGACATCTTCCGCTCTTCCGAGCGG 2100

QY 2149 CTTAAAGAGCTCTCCAGATGCGTCCAGCAATCGACGAGTCTCTGGAGACCAACCGCG 2208  
DB |||||  
QY 2101 CTTAAAGAGCTCTCCAGATGCGTCCAGCAATCGACGAGTCTCTGGAGACCAACCGCG 2160  
DB |||||  
QY 2209 GTAGCCGAGGTGGCTATGGCTTATTCTGTGATCTCTCTGATCCAAATGAAAGCTCTAC 2268  
DB |||||  
QY 2161 GTGCTGAAGTGGCTATGGCTTATTCTGTGATCTCTCTGATCCGAAATGAAAGCTCTAC 2220  
DB |||||  
QY 2269 ACCCTGATTTACTACCTTAAGATGGAGAGGATCGTCAAGTCTGGCGCTCACATCTTG 2328  
DB |||||  
QY 2221 ACCCTGATTTACTACCTTAAGATGGAGAGGATCGTCAAGTCTGGCGCTCACATCTTG 2280  
DB |||||  
QY 2329 GCCATTAAGGATATGGCTCTGCTTCCGCCAGCTGCGGTAAACCAAGTGTGTCAACGCA 2388  
DB |||||  
QY 2281 GCTATTAAGGATATGGCTCTGCTTCCGCCAGCTGCGGTAAACCAAGTGTGTCAACGCA 2340  
DB |||||  
QY 2389 CTGGCCGCTGAATTCGATCTGCCAGTGCAGTGCACACCAAGACACTCGGGTGGCCAG 2448  
DB |||||  
QY 2341 CTGGCCGCTGAATTTGATCTGCCAGTGCAGTGCACACCAAGACACTCGGGTGGCCAG 2400  
DB |||||  
QY 2449 CTGGCAACCTACTTTGCTGAGCTCAAGCTGGTGCAGATGCTGTGACGGTCTTCCGCA 2508  
DB |||||  
QY 2401 CTGGCAACCTACTTTGCTGAGCTCAAGCTGGTGCAGATGCTGTGACGGTCTTCCGCA 2460  
DB |||||  
QY 2509 CCAGTGTCTGGACCACTCTCCAGCCATCCCTGTCTGCCATTTGTCGCTGATTCGGCAC 2568  
DB |||||  
QY 2461 CCAGTGTCTGGACCACTCTCCAGCCATCCCTGTCTGCCATTTGTCGCTGATTCGGCAC 2520  
DB |||||  
QY 2569 ACCCGTCCGATACCGGTTTGGAGCTCGAGGCTGTTTCTGACCTCGAGCGGTACTGGAA 2628  
DB |||||  
QY 2521 ACCCGTCCGATACCGGTTTGGAGCTCGAGGCTGTTTCTGACCTCGAGCGGTACTGGAA 2580  
DB |||||  
QY 2629 GCAGTGGCGGAGTGTACTTGCATCTGAGTCTGGAACCCAGCCCAACCGTCCGCTC 2688  
DB |||||  
QY 2581 GCAGTGGCGGAGTGTACTTGCATCTGAGTCTGGAACCCAGCCCAACCGTCCGCTC 2640  
DB |||||  
QY 2689 TACCGCCACCAAAATCCAGCGGAGTGTCTCAACCTGCTGCACAGGCCACCGCACTG 2748  
DB |||||  
QY 2641 TACCGCCACCAAAATCCAGCGGAGTGTCTCAACCTGCTGCACAGGCCACCGCACTG 2700  
DB |||||  
QY 2749 GGCCTTGGGATCGTTTGCAGTCTCATCGAGACACTACCGACCGGTTAATGAGTCTG 2808  
DB |||||  
QY 2701 GGCCTTGGGATCGTTTGCAGTCTCATCGAGACACTACCGACCGGTTAATGAGTCTG 2760  
DB |||||  
QY 2809 GGAGCCCAACCAAGGTCAACCCATCTCCAGGTTGTTGGCGACCTCGCACTCCACCTC 2868  
DB |||||  
QY 2761 GGAGCCCAACCAAGGTCAACCCATCTCCAGGTTGTTGGCGACCTCGCACTCCACCTC 2820  
DB |||||  
QY 2869 GTTGGTGGGTTGGATCCAGGAGACTTTGCTGCCGATCCCAAAAGTACGACATCCCA 2928  
DB |||||  
QY 2821 GTTGGTGGGTTGGATCCAGGAGACTTTGCTGCCGATCCCAAAAGTACGACATCCCA 2880  
DB |||||  
QY 2929 GACTCTGTATCGGTTCTCTGGCGGCGAGCTTGGTAAACCTCCAGGTGGCTGGCCAGAG 2988  
DB |||||  
QY 2881 GACTCTGTATCGGTTCTCTGGCGGCGAGCTTGGTAAACCTCCAGGTGGCTGGCCAGAG 2940  
DB |||||  
QY 2989 CCACTGGGCAACCGCGCACTGGAAAGCGCTCCGAAGGCAAGGCACTCTGACGGAAGTT 3048  
DB |||||  
QY 2941 CCACTGGGCAACCGCGCACTGGAAAGCGCTCCGAAGGCAAGGCACTCTTGGACGGAAGTT 3000  
DB |||||  
QY 3049 CCTGAGGAAGAGAGGCGGCACTCGAGCTGATGATTCGAAGGAACCTCGCAATAGCTC 3108  
DB |||||  
QY 3001 CCTGAGGAAGAGAGGCGGCACTCGAGCTGATGATTCGAAGGAACCTCGCAAGAGCTC 3060  
DB |||||  
QY 3109 AACCGCTGTGTTCCGGAAGCCAAACGGAAGTTCCTCGAGCACCGTCCGCGCTTCGGC 3168  
DB |||||  
QY 3061 AACCGCTGTGTTCCGGAAGCCAAACGGAAGTTCCTCGAGCACCGTCCGCGCTTCGGC 3120  
DB |||||  
QY 3169 AACCGCTGTGTTCCGGAAGCCAAACGGAAGTTCCTCGAGCACCGTCCGCGCTTCGGC 3228  
DB |||||  
QY 3121 AACCGCTGTGTTCCGGAAGCCAAACGGAAGTTCCTCGAGCACCGTCCGCGCTTCGGC 3180  
DB |||||

QY 3229 TTGATCCGCTGCGAGATGTCGCAACCCCACTCTTGTTCGCTGGATGCGATCTCTGAG 3288  
DB |||||  
QY 3181 TTGATCCGCTGCGAGATGTCGCAACCCCACTCTTGTTCGCTGGATGCGATCTCTGAG 3240  
DB |||||  
QY 3289 CCAGACATTAAGGGTATGCGCAATGTTGGCCAAAGTCAACGCGCAGATCCGCCCAATG 3348  
DB |||||  
QY 3241 CCAGACATTAAGGGTATGCGCAATGTTGGCCAAAGTCAACGCGCAGATCCGCCCAATG 3300  
DB |||||  
QY 3349 CGTGTGCTGACCGCTCCGTTGAGTCTGTCAACGCAACCGCAGAAAGGAGATTCCTCC 3408  
DB |||||  
QY 3301 CGTGTGCTGACCGCTCCGTTGAGTCTGTCAACGCAACCGCAGAAAGGAGATTCCTCC 3360  
DB |||||  
QY 3409 AACAGGCGCATGTTGTGCAACCATTCGCTGGTGTGTCAACCGTCACTGTTGCTGAAGT 3468  
DB |||||  
QY 3361 AACAGGCGCATGTTGTGCAACCATTCGCTGGTGTGTCACTGTGACTGTTGCTGAAGT 3420  
DB |||||  
QY 3469 GATGAGTCAAGGCTGAGATGAGTGCATCATCGAGCTATGAAGATGGAAGCAACA 3528  
DB |||||  
QY 3421 GATGAGTCAAGGCTGAGATGAGTGCATCATCGAGCTATGAAGATGGAAGCAACA 3480  
DB |||||  
QY 3529 ATCACTGCTTCTGTTGACGCAAAATCGATCGGTTGTGTTCTGCTGCAAGAGGTG 3588  
DB |||||  
QY 3481 ATCACTGCTTCTGTTGACGCAAAATCGATCGGTTGTGTTCTGCTGCAAGAGGTG 3540  
DB |||||  
QY 3589 GAAGTGGCGACTTGAATCGTCTGTTCTCTAA 3621  
DB |||||  
QY 3541 GAAGTGGCGACTTGAATCGTCTGTTCTCTAA 3573  
DB |||||  
RESULT 10  
AX120849 3420 bp DNA linear PAT 11-MAY-2001  
LOCUS  
DEFINITION Sequence 765 from Patent EP1108790.  
ACCESSION AX120849  
VERSION AX120849.1 GI:14037564  
KEYWORDS  
SOURCE Corynebacterium glutamicum  
ORGANISM Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriaceae; Corynebacteriinae; Corynebacterium.  
REFERENCE  
1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.,  
Novel polynucleotides  
Patent: EP 1108790-A 765 20-JUN-2001;  
KYOWA HAKKO KOGYO CO., LTD. (JP)  
FEATURES  
1. 3420  
Location/Qualifiers  
/organism="Corynebacterium glutamicum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1718"  
ORIGIN  
Query Match 94.4%; Score 3420; DB 6; Length 3420;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 199 GTGTGACTCACACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGGTAGCAACCGC 258  
DB 1 GTGTGACTCACACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGGTAGCAACCGC 60  
QY 259 GGCAGAAATCGCGGTCCGTCGCTTTCCGTCGAGCACTCGAAACCGGTGCGACCGGTAGCT 318  
DB 61 GGCAGAAATCGCGGTCCGTCGCTTTCCGTCGAGCACTCGAAACCGGTGCGACCGGTAGCT 120  
QY 319 ATTTACCCCGGTGAAGATCGGGATCAATTCACCGCTCTTTTGTCTTGAAGCTTCGGC 378  
DB 121 ATTTACCCCGGTGAAGATCGGGATCAATTCACCGCTCTTTTGTCTTGAAGCTTCGGC 180  
QY 379 ATTTGATCCGAGGCTCACCAAGTCAAGCGGTACCTGGACATCGATGAATATATCGGTGCA 438  
DB 181 ATTTGATCCGAGGCTCACCAAGTCAAGCGGTACCTGGACATCGATGAATATATCGGTGCA 240

QY 439 GCTAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACCGCTTCTGTCTGTAATGCC 498  
Db 241 GCTAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACCGCTTCTGTCTGTAATGCC 300  
QY 499 CAGCTTCCCGGAGTGTGCGGAAACGGCAATCTTTTATTTGGCCCAACCCCAAGAGTT 558  
Db 301 CAGCTTCCCGGAGTGTGCGGAAACGGCAATCTTTTATTTGGCCCAACCCCAAGAGTT 360  
QY 559 CTTGATCTCACCGGTGATAAGTCTGCGCGGTAAACCGCCGGAAGAGGCTGTCTGCCA 618  
Db 361 CTTGATCTCACCGGTGATAAGTCTGCGCGGTAAACCGCCGGAAGAGGCTGTCTGCCA 420  
QY 619 GTTTTGGCGGATCCACCCGAGCAAAACATCGATGAGTCGTTAAAGCGCTGAAGGC 678  
Db 421 GTTTTGGCGGATCCACCCGAGCAAAACATCGATGAGTCGTTAAAGCGCTGAAGGC 480  
QY 679 CAGACTTACCCCATCTTTGTGAAGCAGTGTGCGGTGTGCGGACGCGGTATCGTTTT 738  
Db 481 CAGACTTACCCCATCTTTGTGAAGCAGTGTGCGGTGTGCGGACGCGGTATCGTTTT 540  
QY 739 GTTGTTCACCTGATGAGTTGCGAAATAGCAACAGAGCATCTCGTGAAGCTGAAGCG 798  
Db 541 GTTGTTCACCTGATGAGTTGCGAAATAGCAACAGAGCATCTCGTGAAGCTGAAGCG 600  
QY 799 GCTTTCCGGGATGCGCGGTATATGTGAAACGTGTGTGATTAACCCCTCAGCATATTGAA 858  
Db 601 GCTTTCCGGGATGCGCGGTATATGTGAAACGTGTGTGATTAACCCCTCAGCATATTGAA 660  
QY 859 GTGAGATCTTTGGCGATCACACTGAGAGTGTGACCTTTATGAACGTGACTGTCA 918  
Db 661 GTGAGATCTTTGGCGATCACACTGAGAGTGTGACCTTTATGAACGTGACTGTCA 720  
QY 919 CTGACGCTGCTCACCAAAAGTGTGCGAAATGCGGCAGACAGCATTTGGATCCAGAA 978  
Db 721 CTGACGCTGCTCACCAAAAGTGTGCGAAATGCGGCAGACAGCATTTGGATCCAGAA 780  
QY 979 CTGCGTGTGCAATTTGTCGGATGACGTAAGTTCTGCGCTTCAATTTGGTTACAGGGC 1038  
Db 781 CTGCGTGTGCAATTTGTCGGATGACGTAAGTTCTGCGCTTCAATTTGGTTACAGGGC 840  
QY 1039 GCGGGAACCGTGGAAATCTTGTGCGATGAAAGGCAACACGCTCTCATCGAATGAAC 1098  
Db 841 GCGGGAACCGTGGAAATCTTGTGCGATGAAAGGCAACACGCTCTCATCGAATGAAC 900  
QY 1099 CCACTGATCCAGTTGAGACACACCGGTGACTGAAGAGTCAACGAGGTGACCTGTGTAAG 1158  
Db 901 CCACTGATCCAGTTGAGACACACCGGTGACTGAAGAGTCAACGAGGTGACCTGTGTAAG 960  
QY 1159 GCGCAGATGCGCTGTGCTGTGCAACCTTTGAAGGAATTCGCTGACCCCAAGATAAG 1218  
Db 961 GCGCAGATGCGCTGTGCTGTGCAACCTTTGAAGGAATTCGCTGACCCCAAGATAAG 1020  
QY 1219 ATCAAGACCCACCGGTGACACTGCAAGTCCCGCATCAACCAAGAGATCCAAACACGCG 1278  
Db 1021 ATCAAGACCCACCGGTGACACTGCAAGTCCCGCATCAACCAAGAGATCCAAACACGCG 1080  
QY 1279 TTCGCGCAGATACCGGAATATACCGGTGACCGCTCAACGAGGAGCTGCGCTGCT 1338  
Db 1081 TTCGCGCAGATACCGGAATATACCGGTGACCGCTCAACGAGGAGCTGCGCTGCT 1140  
QY 1339 CTTGACGCTGACGCTCAGCTGCGGTGCGAAATCAACGACACATTTTGACTTCCATGCTGTG 1398  
Db 1141 CTTGACGCTGACGCTCAGCTGCGGTGCGAAATCAACGACACATTTTGACTTCCATGCTGTG 1200  
QY 1399 AAAATGACCTGCGGTGCTGCACTTTGAAATGCTGTTGCTGCTGCAACGCGCGGTG 1458  
Db 1201 AAAATGACCTGCGGTGCTGCACTTTGAAATGCTGTTGCTGCTGCAACGCGCGGTG 1260  
QY 1459 GCTGAGTTACCGGTGCTGCTGTTGCAACCAACATTTGTTGCTGCTGCTGCTGCTG 1518  
Db 1261 GCTGAGTTACCGGTGCTGCTGTTGCAACCAACATTTGTTGCTGCTGCTGCTGCTG 1320  
QY 1519 GAAGAGGACTTCACTTCCAAAGCGCATCGCCACCGGATTTCAATGCCGATCACCCGCACTTC 1578

Db 1321 GAAGAGGACTTCACTTCCAAAGCGCATCGCCACCGGATTCATTTGCCGATCACCCGCACTTC 1380  
QY 1579 CTTTAGGCTTCCACCTGCTGATGATGAGCAGGACGCACTCTTGATTAATTTGGCAGATGTC 1638  
Db 1381 CTTTAGGCTTCCACCTGCTGATGATGAGCAGGACGCACTCTTGATTAATTTGGCAGATGTC 1440  
QY 1639 ACGGTGAAACAGGCTCATGCTGCTGCTCCAAAGGATGTTGAGCTCTCTATCGATGAAGTG 1698  
Db 1441 ACGGTGAAACAGGCTCATGCTGCTGCTCCAAAGGATGTTGAGCTCTCTATCGATGAAGTG 1500  
QY 1699 CTTAAACATCAGAGATCTGCCACTGCCACGCGGTTCCTGAGCCGCTGAAGCAGTTGGC 1758  
Db 1501 CTTAAACATCAGAGATCTGCCACTGCCACGCGGTTCCTGAGCCGCTGAAGCAGTTGGC 1560  
QY 1759 CAGCCGCGTTTGTCTGCTGATCTCGTGAAGCAGCAGCTGCGAGTTACTGATACCAACC 1818  
Db 1561 CAGCCGCGTTTGTCTGCTGATCTCGTGAAGCAGCAGCTGCGAGTTACTGATACCAACC 1620  
QY 1819 TTCCGCGATGCAACACAGTCTTTTGTGCTTGCACCCGAGTCCGCTCATTTGCACTGAAGCT 1878  
Db 1621 TTCCGCGATGCAACACAGTCTTTTGTGCTTGCACCCGAGTCCGCTCATTTGCACTGAAGCT 1680  
QY 1879 GCGCAGAGGCGCTGCAAAAGCTGACTCTGAGCTTTTGTGCTGAGGCTCTGGGCGGCG 1938  
Db 1681 GCGCAGAGGCGCTGCAAAAGCTGACTCTGAGCTTTTGTGCTGAGGCTCTGGGCGGCG 1740  
QY 1939 GCGACCTACGATGTCGCGATGCTTTCCTTTGAGGATCCGTCGAGCAGAGCTCGACGAG 1998  
Db 1741 GCGACCTACGATGTCGCGATGCTTTCCTTTGAGGATCCGTCGAGCAGAGCTCGACGAG 1800  
QY 1999 CTGCGGAGCGGATGCGGAATGTAAACATTCAGATGCTCTTCCGCGCGCGCAACCCGTC 2058  
Db 1801 CTGCGGAGCGGATGCGGAATGTAAACATTCAGATGCTCTTCCGCGCGCGCAACCCGTC 1860  
QY 2059 GGATACACCCGCTACCCAGACTCCGCTCGCGCGCTTTGTTAAGGAAGCTGCCAGCTCC 2118  
Db 1861 GGATACACCCGCTACCCAGACTCCGCTCGCGCGCTTTGTTAAGGAAGCTGCCAGCTCC 1920  
QY 2119 GCGCTGACATCTTCCGCACTTTTCGACGCGCTTAAACAGCTCTTCCAGATGCTGTCAGCA 2178  
Db 1921 GCGCTGACATCTTCCGCACTTTTCGACGCGCTTAAACAGCTCTTCCAGATGCTGTCAGCA 1980  
QY 2179 ATGACGCTGCTGAGACCAACACGCGGTGAGCGAGTGGCTATGCTTATTTCTGGT 2238  
Db 1981 ATGACGCTGCTGAGACCAACACGCGGTGAGCGAGTGGCTATGCTTATTTCTGGT 2040  
QY 2239 GATCTCTCTCATCCAAATGAAAGCTTACACCTCGATTAACCTTAAAGATGGCAGAG 2298  
Db 2041 GATCTCTCTCATCCAAATGAAAGCTTACACCTCGATTAACCTTAAAGATGGCAGAG 2100  
QY 2299 GAGATCTGAGTGTGCGCTCACATCTTGGCCATTAAGGATATGCTGCTGCTGCTG 2358  
Db 2101 GAGATCTGAGTGTGCGCTCACATCTTGGCCATTAAGGATATGCTGCTGCTGCTG 2160  
QY 2359 CCAGCTCGGTAAACCAAGCTGTCACCGCACTGCGCGCTGAATTCGATCTGCCAGTGCAC 2418  
Db 2161 CCAGCTCGGTAAACCAAGCTGTCACCGCACTGCGCGCTGAATTCGATCTGCCAGTGCAC 2220  
QY 2419 GTGCAACCAACGACACTGCGGCTGGCAGCTGGCACTTCTTGTGCTGAGCTCAGCT 2478  
Db 2221 GTGCAACCAACGACACTGCGGCTGGCAGCTGGCACTTCTTGTGCTGAGCTCAGCT 2280  
QY 2479 GGTGAGATGCTGTGAGCGTGTTCGCGACCACTGTCTGGCAGCACTCTCCAGCCATCC 2538  
Db 2281 GGTGAGATGCTGTGAGCGTGTTCGCGACCACTGTCTGGCAGCACTCTCCAGCCATCC 2340  
QY 2539 CTGCTGCAATTTGCTGATTCGCGCACACCGCTGCGATACCGGTTTGGAGCTCGAG 2598  
Db 2341 CTGCTGCAATTTGCTGATTCGCGCACACCGCTGCGATACCGGTTTGGAGCTCGAG 2400  
QY 2599 GCTGTTTCTGACCTCGAGCGCTACTGGAAGCAGTGGCGGAGCTGATCTGCACTTTCAG 2658

Db	2401	GCTGTTTCTTGAACCTCGAGCGGTACTGGGAAGCAGTGGCGGACTGTACCTGCAATTTGAG	2460
Qy	2659	TCCTGGAACCCAGGCCCAACCGGTGCGCTCTACCGCACAGAAATCCAGGCGGACAGTTG	2718
Db	2461	TCCTGGAACCCAGGCCCAACCGGTGCGCTCTACCGCACAGAAATCCAGGCGGACAGTTG	2520
Qy	2719	TCCAACCTCGTGTGACAGAGCCACCGCACTGGGCTTGGCGATCGTTTGGAACTCATCGAA	2778
Db	2521	TCCAACCTCGTGTGACAGAGCCACCGCACTGGGCTTGGCGATCGTTTGGAACTCATCGAA	2580
Qy	2779	GACAACTAGCGAGCGCTTAATGAGATCTGGAGCGCCCAACCAAGTCAACCCATCTCC	2838
Db	2581	GACAACTAGCGAGCGCTTAATGAGATCTGGAGCGCCCAACCAAGTCAACCCATCTCC	2640
Qy	2839	AAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGGGCTGTGGATCCAGCAGACTTT	2898
Db	2641	AAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGGGCTGTGGATCCAGCAGACTTT	2700
Qy	2899	GCTGGCGATCCACAAAGTAGACATCCAGACTCTGTTCATGCGCTTCTGGCGGCGAG	2958
Db	2701	GCTGGCGATCCACAAAGTAGACATCCAGACTCTGTTCATGCGCTTCTGGCGGCGAG	2760
Qy	2959	CTTGGTAACCTCCAGGTGCTGGCCAGAGCCACTGCGCACCGCGCACTGGAGGCGCG	3018
Db	2761	CTTGGTAACCTCCAGGTGCTGGCCAGAGCCACTGCGCACCGCGCACTGGAGGCGCG	2820
Qy	3019	TCCGAGGCGAAGGCACTCTGAGCGAGTCTCTGAGGAGAGAGGCGGCACTCGACGCT	3078
Db	2821	TCCGAGGCGAAGGCACTCTGAGCGAGTCTCTGAGGAGAGAGGCGGCACTCGACGCT	2880
Qy	3079	GATGATTCCAGGAACCTGCAATAGCTCAACCGCTGCTGTTCGCCGAAGCAACCGAA	3138
Db	2881	GATGATTCCAGGAACCTGCAATAGCTCAACCGCTGCTGTTCGCCGAAGCAACCGAA	2940
Qy	3139	GAGTTCTCGAGCACTGTCGCGCTTGGCAACACCTCTGCGCTGATGATCGTGAATTC	3198
Db	2941	GAGTTCTCGAGCACTGTCGCGCTTGGCAACACCTCTGCGCTGATGATCGTGAATTC	3000
Qy	3199	TTCTAGCGCTGTTGGAAGCGCGAGACTTTGATCGCTGCGACATGTGGCCACCCA	3258
Db	3001	TTCTAGCGCTGTTGGAAGCGCGAGACTTTGATCGCTGCGACATGTGGCCACCCA	3060
Qy	3259	CTGCTTGTTCGCTGATGCGATCTCTGAGCCAGAGATAGGGTATGCGCAATGTTGTG	3318
Db	3061	CTGCTTGTTCGCTGATGCGATCTCTGAGCCAGAGATAGGGTATGCGCAATGTTGTG	3120
Qy	3319	GCCAACTCAACGCGCAGATCCGCCCAATCGGTGCGTGAACCGCTCGTGAATCTGTC	3378
Db	3121	GCCAACTCAACGCGCAGATCCGCCCAATCGGTGCGTGAACCGCTCGTGAATCTGTC	3180
Qy	3379	ACCGCAACCGCAGAAAGGAGATTCCTCCAAAGGCGCAATGTTGTCGACCAATTCGCT	3438
Db	3181	ACCGCAACCGCAGAAAGGAGATTCCTCCAAAGGCGCAATGTTGTCGACCAATTCGCT	3240
Qy	3439	GGTGTGTACCGTACCTGTTGCTGAGGTGATGAGTCAAGGCTGAGATGCGATCGCA	3498
Db	3241	GGTGTGTACCGTACCTGTTGCTGAGGTGATGAGTCAAGGCTGAGATGCGATCGCA	3300
Qy	3499	ATCATCGAGGCTATGAAGTAGAGCAACCAATCACTGCTTCTGTGACGCAAAATCGAT	3558
Db	3301	ATCATCGAGGCTATGAAGTAGAGCAACCAATCACTGCTTCTGTGACGCAAAATCGAT	3360
Qy	3559	CGGCTTGTGTTCTGCTGCAACGAGGTGAGGTGGGCACTGATCGTCTGCTGTTTC	3618
Db	3361	CGGCTTGTGTTCTGCTGCAACGAGGTGAGGTGGGCACTGATCGTCTGCTGTTTC	3420

RESULT 11  
BD162966  
LOCUS BD162966 3420 bp DNA linear PAT 17-JAN-2003  
DEFINITION Novel polynucleotide.  
ACCESSION BD162966  
VERSION BD162966.1 Gi:27868728

KEYWORDS	JP 2002191370-A/765.
SOURCE	unidentified
ORGANISM	unidentified
REFERENCE	unclassified.
AUTHORS	1 (bases 1 to 3420)
TITLE	Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
JOURNAL	Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
COMMENT	Novel polynucleotide Patent: JP 2002191370-A 765 09-JUL-2002; KYOWA HAKKO KOGYO CO LTD OS Corynebacterium glutamicum PN JP 2002191370-A/765 PD 09-JUL-2002 PF 15-DEC-2000 JP 2000405096 PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI, PI KEIKO OCHIAI, PI HARUHIKO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO PI OZAKI PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC C12N1/15, PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/ PC 04, C12P13/08, PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC G01N33/566, PC G01N33/569, G01N33/68, G01N37/00, C12P21/08, C12N1/21, C12R1:15), PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15), PC C12N15/00, PC C12N5/00, C12N15/00 CC Novel polynucleotide FH Key Location/Qualifiers FT source 1..3420 FT /organism='Corynebacterium glutamicum'. FEATURES source 1..3420 Location/Qualifiers /organism='unidentified' /mol_type='genomic DNA' /db_xref='taxon:32644'
ORIGIN	Query Match 94.4%; Score 3420; DB 6; Length 3420; Best Local Similarity 100.0%; Pred. No. 0; Matches 3420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	199 GTGTCGATCACACATTTCAACGCTTCCAGCATTTCAAAAGATCTTGGTAGCAACCGC 259
Db	1 GTGTCGATCACACATTTCAACGCTTCCAGCATTTCAAAAGATCTTGGTAGCAACCGC 60
Qy	259 GCGGAAATCGCGTCCGTCGTTCCGTCGACGACTCGAAACCGGTGCGACCGGTAGCT 318
Db	61 GCGGAAATCGCGTCCGTCGTTCCGTCGACGACTCGAAACCGGTGCGACCGGTAGCT 120
Qy	319 ATTTACCCCTGTAAGATCGGGATCATTCACCGCTCTTTTGTCTGAGCTGTCCGC 378
Db	121 ATTTACCCCTGTAAGATCGGGATCATTCACCGCTCTTTTGTCTGAGCTGTCCGC 180
Qy	379 ATTGTCACCGAGCTCACAGTCAAGCGTACCTGACATCGATGAAATTTATCGGTGCA 438
Db	181 ATTGTCACCGAGGCTCACAGTCAAGCGTACCTGACATCGATGAAATTTATCGGTGCA 240
Qy	439 GCTAAAAAGTTAAAGCAGATGCCATTTACCCGGATACCGCTTCTGTCTGAAAAATGCC 498
Db	241 GCTAAAAAGTTAAAGCAGATGCCATTTACCCGGATACCGCTTCTGTCTGAAAAATGCC 300
Qy	499 CAGCTTCCCGGAGTGTGGGAAACGGCATTTATTTATTTGGCCCAACCCAGAGGTT 558
Db	301 CAGCTTCCCGGAGTGTGGGAAACGGCATTTATTTATTTGGCCCAACCCAGAGGTT 360
Qy	559 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAAACCGCGGAGAAAGGCTGGTCTGCCA 618
Db	361 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAAACCGCGGAGAAAGGCTGGTCTGCCA 420
Qy	619 GTTTGGCGGATTCACCCCGAGCAAAACATCATGATGATCGTTAAAGCGCTGAAGGC 678

Db 421 GTTTTGGCGGAATCCACCCCGAGCAAAAATCGATGAGATCGTTTAAAGCGCTGAAGGC 480  
QY CAGACTTACCCCATCTTTGTGAAGCAGTTGCGCGTGTGGCGAGCGCGGTATGCGTTTT 738  
Db 481 CAGACTTACCCCATCTTTGTGAAGCAGTTGCGCGTGTGGCGAGCGCGGTATGCGTTTT 540  
QY GTTGTCTACCTGATGAGCTTCGCAATTTAGCAACAGAGCATCTCGTGAAGCTGAAGCG 798  
Db 541 GTTGTCTACCTGATGAGCTTCGCAATTTAGCAACAGAGCATCTCGTGAAGCTGAAGCG 600  
QY GTTGTCTACCTGATGAGCTTCGCAATTTAGCAACAGAGCATCTCGTGAAGCTGAAGCG 858  
Db 601 GTTGTCTACCTGATGAGCTTCGCAATTTAGCAACAGAGCATCTCGTGAAGCTGAAGCG 660  
QY GTGCAGATCTTTGGCGATCACA CTGGAGAGTTGTACACCTTTATGAACGTGACTCA 918  
Db 661 GTGCAGATCTTTGGCGATCACA CTGGAGAGTTGTACACCTTTATGAACGTGACTCA 720  
QY CTGCAGCGTCTGCACCAAAAGTTGTGAAATTTGGCGAGCAGACAGCATTTGATCCAGAA 978  
Db 721 CTGCAGCGTCTGCACCAAAAGTTGTGAAATTTGGCGAGCAGACAGCATTTGATCCAGAA 780  
QY CTGCAGCGTCTGCACCAAAAGTTGTGAAATTTGGCGAGCAGACAGCATTTGATCCAGAA 1038  
Db 781 CTGCAGCGTCTGCACCAAAAGTTGTGAAATTTGGCGAGCAGACAGCATTTGATCCAGAA 840  
QY GCGGGAACCGTGGAAATCTTTGTGATGAAAGGCAACACAGCTCTCATCGAATGAAC 1098  
Db 841 GCGGGAACCGTGGAAATCTTTGTGATGAAAGGCAACACAGCTCTCATCGAATGAAC 900  
QY CCAAGTATCCAGGTTGAGCACAACCGTGAAGAGTCAACGAGGTGGAACCTGGTGAAG 1158  
Db 901 CCAAGTATCCAGGTTGAGCACAACCGTGAAGAGTCAACGAGGTGGAACCTGGTGAAG 960  
QY GCGGAGATCGGTTGGCTGCTGCTGCAACCTTTGAAGGATTTGGGCTGACCCAGATGAAG 1218  
Db 961 GCGGAGATCGGTTGGCTGCTGCTGCAACCTTTGAAGGATTTGGGCTGACCCAGATGAAG 1020  
QY ATCAAGACCCAGGTTGAGCACAACCGTGAAGAGTCAACGAGGTGGAACCTGGTGAAG 1278  
Db 1021 ATCAAGACCCAGGTTGAGCACAACCGTGAAGAGTCAACGAGGTGGAACCTGGTGAAG 1080  
QY TTTCCGCGCAGATACCGGATATACCGGATACCGGATACCGGATACCGGATACCGGATACCGG 1338  
Db 1081 TTTCCGCGCAGATACCGGATATACCGGATACCGGATACCGGATACCGGATACCGGATACCGG 1140  
QY CTTCAGGTTGAGCACAACCGTGAAGAGTCAACGAGGTGGAACCTGGTGAAG 1398  
Db 1141 CTTCAGGTTGAGCACAACCGTGAAGAGTCAACGAGGTGGAACCTGGTGAAG 1200  
QY AAAATGACCTCGGTTGAGCACAACCGTGAAGAGTCAACGAGGTGGAACCTGGTGAAG 1458  
Db 1201 AAAATGACCTCGGTTGAGCACAACCGTGAAGAGTCAACGAGGTGGAACCTGGTGAAG 1260  
QY GTTGAAGTTCAACCGTGTCTGGTTGCAACCAACATTTGTTTCTGCGTGGTGTCTGGG 1518  
Db 1261 GTTGAAGTTCAACCGTGTCTGGTTGCAACCAACATTTGTTTCTGCGTGGTGTCTGGG 1320  
QY GAAGAGATCTTCACTTCCAGGATACCGGATACCGGATACCGGATACCGGATACCGGATACCGG 1578  
Db 1321 GAAGAGATCTTCACTTCCAGGATACCGGATACCGGATACCGGATACCGGATACCGGATACCGG 1380  
QY CTTTCAGGCTCCAGCTGATGATGAGCAGGAGCAGCATCTCGATTTACTTTGGCAGATGTC 1638  
Db 1381 CTTTCAGGCTCCAGCTGATGATGAGCAGGAGCAGCATCTCGATTTACTTTGGCAGATGTC 1440  
QY ACCGTGAACAGCTCATGTTGAGTTCGCAACAGAGATTTGAGTTCGCAAGCTTCATGATGAAGCTG 1698  
Db 1441 ACCGTGAACAGCTCATGTTGAGTTCGCAACAGAGATTTGAGTTCGCAAGCTTCATGATGAAGCTG 1500  
QY CTTTAACTCAAGATCTGCACTGCAACCGGTTTCCCGTACCGGCTGAAGCAGCTTGGC 1758  
Db 1501 CTTTAACTCAAGATCTGCACTGCAACCGGTTTCCCGTACCGGCTGAAGCAGCTTGGC 1560

QY CCAGCCGCGTTTCTCGTGAATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1818  
Db 1561 CCAGCCGCGTTTCTCGTGAATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1620  
QY TTTCCGCGATGCAACAGCTCTTTGTTGCGACCCGAGTCCGCTCATTTCCGACCTGAAGCCT 1878  
Db 1621 TTTCCGCGATGCAACAGCTCTTTGTTGCGACCCGAGTCCGCTCATTTCCGACCTGAAGCCT 1680  
QY GCGGAGAGCCCGTCCAAAAGCTGACTCTCTGAGCTTTTGTCCGTGAGGCGCTGGGCGCG 1740  
Db 1681 GCGGAGAGCCCGTCCAAAAGCTGACTCTCTGAGCTTTTGTCCGTGAGGCGCTGGGCGCG 1938  
QY GCGGAGAGCCCGTCCAAAAGCTGACTCTCTGAGCTTTTGTCCGTGAGGCGCTGGGCGCG 1938  
Db 1741 GCGGAGAGCCCGTCCAAAAGCTGACTCTCTGAGCTTTTGTCCGTGAGGCGCTGGGCGCG 1800  
QY CTGCGGAGCGGATGCGGAATGTAAACATTTAGAGTCTCTGCGGCGCGCAACACCGTG 2058  
Db 1801 CTGCGGAGCGGATGCGGAATGTAAACATTTAGAGTCTCTGCGGCGCGCAACACCGTG 1860  
QY GGATACACCCCGTACCGAGACTCCGCTGCGCGCGGTTGTTAAAGAGCTGCCACTCC 2118  
Db 2059 GGATACACCCCGTACCGAGACTCCGCTGCGCGCGGTTGTTAAAGAGCTGCCACTCC 2118  
QY GGATACACCCCGTACCGAGACTCCGCTGCGCGCGGTTGTTAAAGAGCTGCCACTCC 1920  
Db 1861 GGATACACCCCGTACCGAGACTCCGCTGCGCGCGGTTGTTAAAGAGCTGCCACTCC 2178  
QY GCGGAGAGCTCTTCCGCTGCTTTCGAGCGGCTTAAACAGCTCTTCCGAGATGCGTCCAGCA 2178  
Db 1921 GCGGAGAGCTCTTCCGCTGCTTTCGAGCGGCTTAAACAGCTCTTCCGAGATGCGTCCAGCA 1980  
QY ATGAGAGCAGTCTTGAGAGCAACACCGGAGTACCGGAGTGGCTATGCTTATTTGCT 2238  
Db 2179 ATGAGAGCAGTCTTGAGAGCAACACCGGAGTACCGGAGTGGCTATGCTTATTTGCT 2238  
QY ATGAGAGCAGTCTTGAGAGCAACACCGGAGTACCGGAGTGGCTATGCTTATTTGCT 2040  
Db 1981 ATGAGAGCAGTCTTGAGAGCAACACCGGAGTACCGGAGTGGCTATGCTTATTTGCT 2298  
QY GATCTCTCTGATCCAAATGAAAGCTCTACCGCTGCTTACTACCTTAAAGATGGCAGAG 2298  
Db 2041 GATCTCTCTGATCCAAATGAAAGCTCTACCGCTGCTTACTACCTTAAAGATGGCAGAG 2100  
QY GAGATGCTTAAAGTCTGCGGCTCACTCTTGGCGCTTAAAGATGCTGCTGCTTCCG 2358  
Db 2299 GAGATGCTTAAAGTCTGCGGCTCACTCTTGGCGCTTAAAGATGCTGCTGCTTCCG 2358  
QY GAGATGCTTAAAGTCTGCGGCTCACTCTTGGCGCTTAAAGATGCTGCTGCTTCCG 2160  
Db 2101 GAGATGCTTAAAGTCTGCGGCTCACTCTTGGCGCTTAAAGATGCTGCTGCTTCCG 2418  
QY CCAGCTGCGGTAAACAGCTGCTCACCGCTGCGCGCTGGAATTCGATCTGCCAGTGCAC 2418  
Db 2359 CCAGCTGCGGTAAACAGCTGCTCACCGCTGCGCGCTGGAATTCGATCTGCCAGTGCAC 2220  
QY CCAGCTGCGGTAAACAGCTGCTCACCGCTGCGCGCTGGAATTCGATCTGCCAGTGCAC 2478  
Db 2161 CCAGCTGCGGTAAACAGCTGCTCACCGCTGCGCGCTGGAATTCGATCTGCCAGTGCAC 2478  
QY GTGCACACCCAGCACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 2280  
Db 2419 GTGCACACCCAGCACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 2478  
QY GTGCAGAGTCTGCTGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2538  
Db 2221 GTGCAGAGTCTGCTGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2538  
QY GTGCAGAGTCTGCTGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2340  
Db 2479 GTGCAGAGTCTGCTGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2538  
QY GTGCAGAGTCTGCTGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2460  
Db 2281 GTGCAGAGTCTGCTGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2538  
QY GTGCAGAGTCTGCTGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2538  
Db 2539 GTGCAGAGTCTGCTGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2340  
QY GTGCAGAGTCTGCTGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2538  
Db 2341 GTGCAGAGTCTGCTGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2400  
QY GCTGTTTCTGACCTCGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2658  
Db 2599 GCTGTTTCTGACCTCGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2658  
QY GCTGTTTCTGACCTCGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2460  
Db 2401 GCTGTTTCTGACCTCGAGCGGCTTCCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2538  
QY TCTGGAACCCAGCGGCTTCCGCGCTGCTGCGGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2718  
Db 2659 TCTGGAACCCAGCGGCTTCCGCGCTGCTGCGGCTGCTGCGGCTGCGGCTGCGGCTGCGGCT 2520  
QY TCTGGAACCCAGCGGCTTCCGCGCTGCTGCGGCTGCTGCGGCTGCGGCTGCGGCTTGG 2778  
Db 2461 TCTGGAACCCAGCGGCTTCCGCGCTGCTGCGGCTGCTGCGGCTGCGGCTTGG 2778  
QY TCCAACTGCTGCTGACAGCGGCTTCCGCGCTGCTGCGGCTTGGTTCGAACTCCTCGAA 2580  
Db 2719 TCCAACTGCTGCTGACAGCGGCTTCCGCGCTGCTGCGGCTTGGTTCGAACTCCTCGAA 2580  
QY GACAACTGCTGCTGACAGCGGCTTCCGCGCTGCTGCGGCTTGGTTCGAACTCCTCGAA 2838  
Db 2779 GACAACTGCTGCTGACAGCGGCTTCCGCGCTGCTGCGGCTTGGTTCGAACTCCTCGAA 2838  
QY GACAACTGCTGCTGACAGCGGCTTCCGCGCTGCTGCGGCTTGGTTCGAACTCCTCGAA 2640  
Db 2581 GACAACTGCTGCTGACAGCGGCTTCCGCGCTGCTGCGGCTTGGTTCGAACTCCTCGAA 2640

QY	2839	AAGGTTGTTGGCGACCTCGACCTCCACCTCGTTGGTGGCGGTGTGATCCAGCAGACTTT	2899		
Db	2641	AAGGTTGTTGGCGACCTCGACCTCCACCTCGTTGGTGGCGGTGTGATCCAGCAGACTTT	2700		
QY	2899	GCTGCCGATCCACAAAGATGACGATCCCAAGATCTGTGTCAATCCGCTTCTCTGCGCGCGAG	2958		
Db	2701	GCTGCCGATCCACAAAGATGACGATCCCAAGATCTGTGTCAATCCGCTTCTCTGCGCGCGAG	2760		
QY	2959	CTTGGTAAACCTTCAGTGGCTGGCGAGGACCACTGGCGACCCGCGCACTGGAAGCGCG	3018		
Db	2761	CTTGGTAAACCTTCAGTGGCTGGCGAGGACCACTGGCGACCCGCGCACTGGAAGCGCG	2820		
QY	3019	TCCGAAGCGCAAGCCTCTGACGGAAGTCTCTGAGGAAGAGCAGCGCACTCGACGCT	3078		
Db	2821	TCCGAAGCGCAAGCCTCTGACGGAAGTCTCTGAGGAAGAGCAGCGCACTCGACGCT	2880		
QY	3079	GATGATTCCAGGAAGCTGCGCAATAGCTCAACCGCTGCTGTTCCTCCGAGCGCAACCGAA	3138		
Db	2881	GATGATTCCAGGAAGCTGCGCAATAGCTCAACCGCTGCTGTTCCTCCGAGCGCAACCGAA	2940		
QY	3139	GAGTTCCTCGAGCAACGCTCGCGCTTCGGCAACACCTCTCGCTGGATGATCGTGAATTC	3198		
Db	2941	GAGTTCCTCGAGCAACGCTCGCGCTTCGGCAACACCTCTCGCTGGATGATCGTGAATTC	3000		
QY	3199	TTCTACGGCTGCTGGAAGCGCGGAGCTTTCATCCGCTCGCTCGAGATGCGCACCCCA	3258		
Db	3001	TTCTACGGCTGCTGGAAGCGCGGAGCTTTCATCCGCTCGCTCGAGATGCGCACCCCA	3060		
QY	3259	CTGCTTGTTCGCTGGAATGCGATCTCTGAGCCAGACGATAAAGGTATGCGCAATGTTGTG	3318		
Db	3061	CTGCTTGTTCGCTGGAATGCGATCTCTGAGCCAGACGATAAAGGTATGCGCAATGTTGTG	3120		
QY	3319	GCCAACTGCAAGCGGAGATCGGCCCAATCGCTGTGGTGACCGCTCGGTGAGTCTGTC	3378		
Db	3121	GCCAACTGCAAGCGGAGATCGGCCCAATCGCTGTGGTGACCGCTCGGTGAGTCTGTC	3180		
QY	3379	ACCGCAACCGCAGAAAAGGAGATTCCTCCAAAGGGCCATGTTGCTGCACCATTCGCT	3438		
Db	3181	ACCGCAACCGCAGAAAAGGAGATTCCTCCAAAGGGCCATGTTGCTGCACCATTCGCT	3240		
QY	3439	GGTGTGTCACCGTGAATGTTGCTGAAGGTGATGAGTCAAGCTCGAGATGAGTCGCA	3498		
Db	3241	GGTGTGTCACCGTGAATGTTGCTGAAGGTGATGAGTCAAGCTCGAGATGAGTCGCA	3300		
QY	3499	ATCATCGAGCTATGAAGTGAAGCAACATCACTGCTTCTGTGTCAGCGCAAAATCGAT	3558		
Db	3301	ATCATCGAGCTATGAAGTGAAGCAACATCACTGCTTCTGTGTCAGCGCAAAATCGAT	3360		
QY	3559	CGGTTGTGTTCTGCTGCAACGAAGGTGAAGGTGCGCACTTGATCGTCTGCTTCC	3618		
Db	3361	CGGTTGTGTTCTGCTGCAACGAAGGTGAAGGTGCGCACTTGATCGTCTGCTTCC	3420		
RESULT 12					
AX453601					
LOCUS					
DEFINITION					
Sequence 1 from Patent WO0231158.					
ACCESSION					
AX453601					
VERSION					
AX453601.1 GI:21712841					
KEYWORDS					
SOURCE					
Corynebacterium glutamicum					
ORGANISM					
Corynebacterium glutamicum					
REFERENCE					
1					
AUTHORS					
Hanke, P.D.					
TITLE					
Feedback-resistant pyruvate carboxylase gene from corynebacterium					
JOURNAL					
Patent: WO 0231158-A 1 18-APR-2002;					
ARCHER-DANIELS-MIDLAND COMPANY (US)					
LOCATION/Qualifiers					
1. .3474					
/organism="Corynebacterium glutamicum"					
/mol_type="unassigned DNA"					

CDS	/db xref="taxon:1718"				
	1. _3474				
	/note="unnamed protein product"				
	/codon_start=1				
	/transl_table=11				
	/protein_id="CAD38136.1"				
	/db xref="GI:21712842"				
	/db xref="RENTREMBL:CAD38136"				
	/translation="MTAITLGGLLKIIITLVSTHTSSTLPAPFKILVANRGEIAVRA FRALENGAATVAIYPREDRGSFHRSPFAGEAVRITGEGSPVKAYLIDIDEIGAAKVK ADAIYGVGFLSENAOLARECAENGITFLGTPDEVLDLTDGKSRVTAAKAGLPLVLA ESTPSKNIDIVKSAQGTQYPIFKVAGGGGGMFVSAPDELKLAKEASEAEAA FGSGVTERAVINPQHIEVQIIGDRTGEVHLYYERDCLQRHKKVLEIAPQHLDLP ELDRICADAVKFCRIGYAGTGVFLVDEKGNHVFIEWNPRIQHTVHTTITVRSFG LVKQMLAAGATLKEIGTQDKIKTHGAALQCRITTEDPNFNGFRITGTTIYRSPG GAGVRAQALGGEITAFHDSMLVMTCRGSDFTAVARAQALAEFTVSGVATIG FLALIREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLAADVTKNKGHPK DVAAPDKNIKDLPLGRSRDLKQGPAAFDRLREODALAVTDITTFPRDAHQSLL ATVRSFALPAAEAVALKTPILLSVEANGGATVDVAMRFLPEDPDLDELREAMEN VNTQMLRGNTVGYTPYPSVCAAFKENASGVDI FRI PDALNDVSOVRPAIDAVL ETNTAVAEVAMAYSDLSDENEKLTIDYIKVAEIEIVKSGAHILAIKDWAGLLRPA VTKLVTALREFDLPVHVDHTDAGGOLATYFAAGAGADAVDAGAPLSTGTSQSL SAIVAAFAHTRDRTGLEAVSDLEPYWEAVRGLYLPFESGTPGTGVRVHVGDPQ LSNLRAQATLGLADRFELIEDNYAAVNEMLRPTKVTSPSSKVVGLALHLVAGVDP ADPAADPKYDIIPDSVIAFLRGLGNPFGWPEPLTRALEGRSEKAPLTVPEBEQ AHLDAADS KERRNSLNEILLFPKTEBFLERHRRFGNTSALDDREFPVGLVEGRTELIR LPVTRPLLVRDLALSEPDDKGMNVVAVNGQIRPMRVDRSVESVTAARAKDSN KGVHAPFAGVTVTVAGDEVKAGDAVAIIENKMEATITASVDGKIERVVVPAATK VEGDGLIVVVS"				
ORIGIN	Query Match	93.9%; Score 3398.8; DB 6; Length 3474;			
	Best Local Similarity	98.6%; Pred. No. 0;			
	Matches 3427; Conservative	0; Mismatches 47; Indels 0; Gaps 0;			
QY	148	GTGACTGCTATCACCCCTTGGCGTCTCTTGTGTAAGGAATAATTAATCTAGTGTGCACT	207		
Db	1	GTGACTGCTATCACCCCTTGGCGTCTCTTGTGTAAGGAATAATTAATCTAGTGTGCACT	60		
QY	208	CACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGTAGCAACCCGCGGGAATC	267		
Db	61	CACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGTAGCAACCCGCGGGAATC	120		
QY	268	CGGTCGCTGCTTTCCGTGCGAGCACTCGAAAACCGGTGCGACCGGTAGCTATTATACCC	327		
Db	121	CGGTCGCTGCTTTCCGTGCGAGCACTCGAAAACCGGTGCGACCGGTAGCTATTATACCC	180		
QY	328	CGTGAAGATCGGGATCATTCACCGCTCTTTGCTTCTGAAGCTGTCGCGATTTGTTACC	387		
Db	181	CGTGAAGATCGGGATCATTCACCGCTCTTTGCTTCTGAAGCTGTCGCGATTTGTTACC	240		
QY	388	GAAAGCTCACAGTCAAGGCGTACCTGGACATCGATGAAATTTATCGGTGCGAGCTTAA	447		
Db	241	GAAAGCTCACAGTCAAGGCGTACCTGGACATCGATGAAATTTATCGGTGCGAGCTTAA	300		
QY	448	GTTAAGCAGATGCCATTTACCCGGATACGGCTTCTGCTGTAATGCCAGCTTGC	507		
Db	301	GTTAAGCAGATGCCATTTACCCGGATACGGCTTCTGCTGTAATGCCAGCTTGC	360		
QY	508	CGCGAGTGTGCGAAAACCGCATTAATTTATTCGCCCAACCCCGAGGTTCTTGTATCTC	567		
Db	361	CGCGAGTGTGCGAAAACCGCATTAATTTATTCGCCCAACCCCGAGGTTCTTGTATCTC	420		
QY	568	ACCGGTGATGATCTCGCGCGTAAACCCCGGAGAGGCTGCTGCCAGTTTTCGCG	627		
Db	421	ACCGGTGATGATCTCGCGCGTAAACCCCGGAGAGGCTGCTGCCAGTTTTCGCG	480		
QY	628	GAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAGCGCTGAAGCGCAGACTTAC	687		
Db	481	GAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAGCGCTGAAGCGCAGACTTAC	540		
QY	688	CCCATCTTTGTGAAGCGATTTGCCGCTGCTGGCGACCGCGTATCGCTTTTGTGCTTCA	747		



Db 541 CCATCTTTGTAAGGCAATGTCGGTGGTGGCGGACCGCGTATGCGCTTTGTTCTTCA 600  
Qy 748 CTTGATGAGCTTCGCAATATAGCAACAGAGCAATCTCGTGAAGCTGAAGCGGCTTTGGC 807  
Db 601 CTTGATGAGCTTCGCAATATAGCAACAGAGCAATCTCGTGAAGCTGAAGCGGCTTTGGC 660  
Qy 808 GATGCGCGGTATATGTCGAACGCTGTGATTAACCTCAGCATATTTGAAGTGCAGATC 867  
Db 661 GACGGTTCGGTATATGTCGAACGCTGTGATTAACCTCAGCATATTTGAAGTGCAGATC 720  
Qy 868 CTTGGCGATCACTGAGAGAGTTGTACACCTTTATGAAGTGAAGTGCAGATC 927  
Db 721 CTTGGCGATCACTGAGAGAGTTGTACACCTTTATGAAGTGAAGTGCAGATC 780  
Qy 928 CGTCACCAAAAGTGTGCAAAATTCGCGCAGCAGCATTTGGATCCAGAACTGCGGTAT 987  
Db 781 CGTCACCAAAAGTGTGCAAAATTCGCGCAGCAGCATTTGGATCCAGAACTGCGGTAT 840  
Qy 988 CGCATTTTGTGGGATGCAAGTAAAGTTCTGCGCTCCATTTGATACCAAGGCGCGGAAAC 1047  
Db 841 CGCATTTTGTGGGATGCAAGTAAAGTTCTGCGCTCCATTTGATACCAAGGCGCGGAAAC 900  
Qy 1048 GTGGAATTTGTGGGATGCAAAAGGCAACCAAGTCTTCTATCGAAATGAACCCAGATC 1107  
Db 901 GTGGAATTTGTGGGATGCAAAAGGCAACCAAGTCTTCTATCGAAATGAACCCAGATC 960  
Qy 1108 CAGGTTGAGCACAACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1167  
Db 961 CAGGTTGAGCACAACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020  
Qy 1168 CGTTGGCTGTGTCGAACCTTGAAGGAATTTGGTCTGACCCAGATGAAGTGAAGTGAAGT 1227  
Db 1021 CGTTGGCTGTGTCGAACCTTGAAGGAATTTGGTCTGACCCAGATGAAGTGAAGTGAAGT 1080  
Qy 1228 CACGCTGAGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1287  
Db 1081 CACGCTGAGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1140  
Qy 1288 GATACCGGAATATCACCGGTATCGCTCACAGGCGAGTGGCTGTGCTGAGCGT 1347  
Db 1141 GATACCGGAATATCACCGGTATCGCTCACAGGCGAGTGGCTGTGCTGAGCGT 1200  
Qy 1348 GCAGCTGAGCTCGGTGGGGAATCACCGCAGTCTTGAAGTGAAGTGAAGTGAAGTGAAGT 1407  
Db 1201 GCAGCTGAGCTCGGTGGGGAATCACCGCAGTCTTGAAGTGAAGTGAAGTGAAGTGAAGT 1260  
Qy 1408 TGGCTGTGTCGAATTTGAAGTGTGCTGTGTCAGCAGCGCGGTGGCTGAGTTC 1467  
Db 1261 TGGCTGTGTCGAATTTGAAGTGTGCTGTGTCAGCAGCGCGGTGGCTGAGTTC 1320  
Qy 1468 ACCGTGTGCTGTGTCGAACCAATTTGCTGTGTCAGTCTTGAAGTGAAGTGAAGTGAAGT 1527  
Db 1321 ACCGTGTGCTGTGTCGAACCAATTTGCTGTGTCAGTCTTGAAGTGAAGTGAAGTGAAGT 1380  
Qy 1528 TTCCTTCAAGCGCATCGCCACCGATTCATTCGCGATCACCGCATCCCTCTTCAAGCT 1587  
Db 1381 TTCCTTCAAGCGCATCGCCACCGATTCATTCGCGATCACCGCATCCCTCTTCAAGCT 1440  
Qy 1588 CCACCTGTGATGATGAGCAGGAGCGATCTGGAATTAATTCGAGATGTCACCGTGAAC 1647  
Db 1441 CCACCTGTGATGATGAGCAGGAGCGATCTGGAATTAATTCGAGATGTCACCGTGAAC 1500  
Qy 1648 AAGCCTCATGTGTGCTGCCAAAGGATGTCAGCTCCATTCGATGAAGTGCCTTAACATC 1707  
Db 1501 AAGCCTCATGTGTGCTGCCAAAGGATGTCAGCTCCATTCGATGAAGTGCCTTAACATC 1560  
Qy 1708 AAGGATCTGCCACTGCCAGCGGTTCCGTGACCGCTGAAGCAGCTTGGCCCGCGCG 1767  
Db 1561 AAGGATCTGCCACTGCCAGCGGTTCCGTGACCGCTGAAGCAGCTTGGCCCGCGCG 1620  
Qy 1768 TTTGCTCGTGAATCTCGTGAGCAGCAGCAGTGGCAGTGAATGATACCACTTCCGCGAT 1827  
Db 1621 TTTGCTCGTGAATCTCGTGAGCAGCAGCAGTGGCAGTGAATGATACCACTTCCGCGAT 1680

Qy 1828 GCACACAGTCTTTGCTTGGACCGGAGTGGCTCATTCGACCTGAAGCCTGCGGCGAG 1887  
Db 1681 GCACACAGTCTTTGCTTGGACCGGAGTGGCTCATTCGACCTGAAGCCTGCGGCGAG 1740  
Qy 1888 GCGGTCGCAAAAGCTGACTCTCTGAGCTTTGTCGAGGCTTGGGCGGCGGCGACTTAC 1947  
Db 1741 GCGGTCGCAAAAGCTGACTCTCTGAGCTTTGTCGAGGCTTGGGCGGCGGCGACTTAC 1800  
Qy 1948 GATGTTGGCGATGCGTTCTCTTTGAGGATCGGTGGGACAGCTCGAGCTGCGGAG 2007  
Db 1801 GATGTTGGCGATGCGTTCTCTTTGAGGATCGGTGGGACAGCTCGAGCTGCGGAG 1860  
Qy 2008 GCGATGCGCAATGTAACATTCAGATGCTGCTTGGCGCGCAACACCGTGGGATACAC 2067  
Db 1861 GCGATGCGCAATGTAACATTCAGATGCTGCTTGGCGCGCAACACCGTGGGATACAC 1920  
Qy 2068 CGGTACCCAGACTCGCTGTCGGCGGCTTTGTTAAGGAGCTGCGAGCTCGGCGTGGAC 2127  
Db 1921 CGGTACCCAGACTCGCTGTCGGCGGCTTTGTTAAGGAGCTGCGAGCTCGGCGTGGAC 1980  
Qy 2128 ATCTTCCGATCTTCCAGCGCTTAAACGACCTCTCCAGATGCGCTCCAGCAATCGAC 2187  
Db 1981 ATCTTCCGATCTTCCAGCGCTTAAACGACCTCTCCAGATGCGCTCCAGCAATCGAC 2040  
Qy 2188 GTCTTGGAGCAACACCGCGGTAGCGGTGCTTATGCTTATTTCTGCTGATCTCTCT 2247  
Db 2041 GTCTTGGAGCAACACCGCGGTAGCGGTGCTTATGCTTATTTCTGCTGATCTCTCT 2100  
Qy 2248 GATCCAAATGAAGCTCTTACACCTGCTTAAACGAGTGGAGAGGAGATCGTC 2307  
Db 2101 GATCCAAATGAAGCTCTTACACCTGCTTAAACGAGTGGAGAGGAGATCGTC 2160  
Qy 2308 AAGTCTGGCGTCAATCTTGGCCATTAAGGATGCTGCTGCTTCTGCTGCGGCGAGTGG 2367  
Db 2161 AAGTCTGGCGTCAATCTTGGCCATTAAGGATGCTGCTGCTTCTGCTGCGGCGAGTGG 2220  
Qy 2368 GTAAACCAAGCTGCTCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2427  
Db 2221 GTAAACCAAGCTGCTCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
Qy 2428 CACGACACTGCGGCTGGCGAGCTGCAACCTTCTTGTGCTGAGCTCAAGCTGGTGCAGAT 2487  
Db 2281 CACGACACTGCGGCTGGCGAGCTGCAACCTTCTTGTGCTGAGCTCAAGCTGGTGCAGAT 2340  
Qy 2488 GCTGTTGAGGCTGCTTCCGACCACTGCTGCGACCACTCCAGCCATCCCTGCTGCTGCC 2547  
Db 2341 GCTGTTGAGGCTGCTTCCGACCACTGCTGCGACCACTCCAGCCATCCCTGCTGCTGCC 2400  
Qy 2548 ATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2607  
Db 2401 ATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460  
Qy 2608 GACCTGAGCGCTTCTGCGGAGCTGCGGAGCTGTACTGCTGCTGCTGCTGCTGCTGCTGCT 2667  
Db 2461 GACCTGAGCGCTTCTGCGGAGCTGCGGAGCTGTACTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
Qy 2668 CAGGCGCCAAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2727  
Db 2521 CAGGCGCCAAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580  
Qy 2728 CGTGACAGGCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2787  
Db 2581 CGTGACAGGCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640  
Qy 2788 GCAGCGCTTAATGAGATGCTGGGACCGCCAAACCAAGGTCACCCCATCTCTCCAAAGTGT 2847  
Db 2641 GCAGCGCTTAATGAGATGCTGGGACCGCCAAACCAAGGTCACCCCATCTCTCCAAAGTGT 2700  
Qy 2848 GCGGACCTCGGACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2907  
Db 2701 GCGGACCTCGGACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760



[illegible]

Query Match 93.9%; Score 3398.8; DB 6; Length 3474;

Db 1021 CGCTTGGCTGCTGGTCAACCTTTGAAGGAATTGGGCTCTGACCCAAAGATAAGATCAAGACC 1080  
QY 1228 CAGGCTGACGACTCGAGTGGCGCATCACCGAGAGATCAAAACAGCGCTTCGGGCCA 1287  
Db 1081 CACGCTGACGACTCGAGTGGCGCATCACCGAGAGATCAAAACAGCGCTTCGGGCCA 1140  
QY 1288 GATACCGGAATCATCACCGCGTACCGCTCACCGAGCGAGCTGGCGTTGCTTTGAGCGGT 1347  
Db 1141 GATACCGGAATCATCACCGCGTACCGCTCACCGAGCGAGCTGGCGTTGCTTTGAGCGGT 1200  
QY 1348 GAGCTCAGCTCGGTGGGGAATACACCGACACTTTGACTCATGCTCGGTGGAATGACC 1407  
Db 1201 GAGCTCAGCTCGGTGGGGAATACACCGACACTTTGACTCATGCTCGGTGGAATGACC 1260  
QY 1408 TGGCTGGTTCCGACTTTGAAACTGCTGTGCTCGTGACAGCGCGCTGGCTGAGTTC 1467  
Db 1261 TGGCTGGTTCCGACTTTGAAACTGCTGTGCTCGTGACAGCGCGCTGGCTGAGTTC 1320  
QY 1468 ACCGTGCTGGTGGTGAACCAACATTTGGTTCTTGGCGTGGCTGCTCGGGGAAGAGAC 1527  
Db 1321 ACCGTGCTGGTGGTGAACCAACATTTGGTTCTTGGCGTGGCTGCTCGGGGAAGAGAC 1380  
QY 1528 TTCACTTCCAGCGCATCGCCACCGGATTCATTGGCGGATCACCGCACCTCTCTTCAGGCT 1587  
Db 1381 TTCACTTCCAGCGCATCGCCACCGGATTCATTGGCGGATCACCGCACCTCTCTTCAGGCT 1440  
QY 1588 CCACCTGTGATGATGAGCAGGAGCGCATCTCTGGATTAATTGGCAGATGTCAACGTGAAC 1647  
Db 1441 CCACCTGTGATGATGAGCAGGAGCGCATCTCTGGATTAATTGGCAGATGTCAACGTGAAC 1500  
QY 1648 AAGCTCATGGTGGCTGCAAGGATGTTGACAGTCTATTCGATAGCTGCTCAATCAATC 1707  
Db 1501 AAGCTCATGGTGGCTGCAAGGATGTTGACAGTCTATTCGATAGCTGCTCAATCAATC 1560  
QY 1708 AAGGATCTGCCACTGCCACCGGTTCCCGTGACCGCTTGAAGCAGCTTTGGCCAGCGCG 1767  
Db 1561 AAGGATCTGCCACTGCCACCGGTTCCCGTGACCGCTTGAAGCAGCTTTGGCCAGCGCG 1620  
QY 1768 TTTGCTCGTGAATCTCGTGAGCAGACGCACTGGCAGTACTGATACCACTTCGGGAT 1827  
Db 1621 TTTGCTCGTGAATCTCGTGAGCAGACGCACTGGCAGTACTGATACCACTTCGGGAT 1680  
QY 1828 GCACACCAAGCTTTGCTTGGCAGCCGAGTCCGCTCATTTGCACTGAAGCCTCGCGCAGAG 1887  
Db 1681 GCACACCAAGCTTTGCTTGGCAGCCGAGTCCGCTCATTTGCACTGAAGCCTCGCGCAGAG 1740  
QY 1888 GCGTGCAGCAAGCTGACTCTGAGCTTTTTCCTGAGGAGCTGGGGCGGCGGACCTAC 1947  
Db 1741 GCGTGCAGCAAGCTGACTCTGAGCTTTTTCCTGAGGAGCTGGGGCGGCGGACCTAC 1800  
QY 1948 GATGTGGCGATCGCTTCTCTTTGAGGATCGTGGGACAGGCTCGACGAGCTGCGGAG 2007  
Db 1801 GATGTGGCGATCGCTTCTCTTTGAGGATCGTGGGACAGGCTCGACGAGCTGCGGAG 1860  
QY 2008 GCGATCGCGAATGTAACAATTGAGATGCTGCTTGGCGGCGGCAACACCGTGGGATACAC 2067  
Db 1861 GCGATCGCGAATGTAACAATTGAGATGCTGCTTGGCGGCGGCAACACCGTGGGATACAC 1920  
QY 2068 CGGTACCCAGACTCCGCTGCGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 2127  
Db 1921 CGGTACCCAGACTCCGCTGCGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 1980  
QY 2128 ATCTTCCGCACTTCGACGCGCTTAAACGAGCTCTCCAGATGCGTCCAGCAATCGAGCA 2187  
Db 1981 ATCTTCCGCACTTCGACGCGCTTAAACGAGCTCTCCAGATGCGTCCAGCAATCGAGCA 2040  
QY 2188 GTCTTGGAGACCAACACCGCGTAGCCGAGTGGCTATGCTTATCTGGTGATCTCTCT 2247  
Db 2041 GTCTTGGAGACCAACACCGCGTAGCCGAGTGGCTATGCTTATCTGGTGATCTCTCT 2100  
QY 2248 GATCCAAATGAAGCTTACACCTCGGATTAACCTTAAAGATGGCAGAGAGATCGTC 2307  
Db 2101 GATCCAAATGAAGCTTACACCTCGGATTAACCTTAAAGATGGCAGAGAGATCGTC 2160

QY 2308 AAGTCTGGCGCTCAATCTTGGCCATTAAAGGATATGGCTGGTCTGCTTGGCCCGAGTCGG 2367  
Db 2161 AAGTCTGGCGCTCAATCTTGGCCATTAAAGGATATGGCTGGTCTGCTTGGCCCGAGTCGG 2220  
QY 2368 GTAACCAAGCTGCTCAACCGCACTGCGCGTGAATTCGATCTGCCAGTGACGTGCACACC 2427  
Db 2221 GTAACCAAGCTGCTCAACCGCACTGCGCGTGAATTCGATCTGCCAGTGACGTGCACACC 2280  
QY 2428 CACGACACTCGCGTGGCGAGCTGGCAACTTCTTGTGTCAGCTCAAGCTGTCGAGAT 2487  
Db 2281 CACGACACTCGCGTGGCGAGCTGGCAACTTCTTGTGTCAGCTCAAGCTGTCGAGAT 2340  
QY 2488 GCTGTTGAGCGTGGCTTCGCGCACTGCTGGCACCACTCCAGCACTCCCTGCTGTCGCC 2547  
Db 2341 GCTGTTGAGCGTGGCTTCGCGCACTGCTGGCACCACTCCAGCACTCCCTGCTGTCGCC 2400  
QY 2548 ATTGTTGCTGCATTCGCGCACACCGCTCGCGATACCGGTTTGAAGCTCGAGCTGTTCT 2607  
Db 2401 ATTGTTGCTGCATTCGCGCACACCGCTCGCGATACCGGTTTGAAGCTCGAGCTGTTCT 2460  
QY 2608 GACCTGAGCGCTACTGCGGAAGCAGTGGCGGACTGTACTCTGCCATTTGAGTCTGGAACC 2667  
Db 2461 GACCTGAGCGCTACTGCGGAAGCAGTGGCGGACTGTACTCTGCCATTTGAGTCTGGAACC 2520  
QY 2668 CAGGCGCCAAACCGTTCGCGTCTACCGCCACGAAATCCAGGCGGACAGTGTGTCACACCTG 2727  
Db 2521 CAGGCGCCAAACCGTTCGCGTCTACCGCCACGAAATCCAGGCGGACAGTGTGTCACACCTG 2580  
QY 2728 CGTGACAGCGCACCGCACTGGGCTTGGGATCGTTCGAACTCATCGAAGCAACTAC 2787  
Db 2581 CGTGACAGCGCACCGCACTGGGCTTGGGATCGTTCGAACTCATCGAAGCAACTAC 2640  
QY 2788 GCAGCGCTTAATGAGATGTGGAGCGCCCAACAAAGTCAACCCATCTTCAAGGTTGT 2847  
Db 2641 GCAGCGCTTAATGAGATGTGGAGCGCCCAACAAAGTCAACCCATCTTCAAGGTTGT 2700  
QY 2848 GCGGACCTCGCACTCCACCTCGTGGTGGGCTGGATCCAGCAGACTTTGCTGCCGAT 2907  
Db 2701 GCGGACCTCGCACTCCACCTCGTGGTGGGCTGGATCCAGCAGACTTTGCTGCCGAT 2760  
QY 2908 CCAAAAGTACGACATCCAGACTCTGTCTCATCGGTTCTTGGCGGCGAGCTTGGTAAC 2967  
Db 2761 CCAAAAGTACGACATCCAGACTCTGTCTCATCGGTTCTTGGCGGCGAGCTTGGTAAC 2820  
QY 2968 CTTCCAGGTGGCTGGCAGCAGCACTGGCACCCGCGCACTGGAGGCGCTCCGAGGC 3027  
Db 2821 CTTCCAGGTGGCTGGCAGCAGCACTGGCACCCGCGCACTGGAGGCGCTCCGAGGC 2880  
QY 3028 AAGGCACTCTGACGGAAGTTCCTGAGGAAGAGAGCGGCACTCGACGCTGATGATCC 3087  
Db 2881 AAGGCACTCTGACGGAAGTTCCTGAGGAAGAGAGCGGCACTCGACGCTGATGATCC 2940  
QY 3088 AAGGAACGTGCMATAGCTTCAACCGCTGCTTCCCGAAGCCCAACCGAAGTTCCTC 3147  
Db 2941 AAGGAACGTGCMATAGCTTCAACCGCTGCTTCCCGAAGCCCAACCGAAGTTCCTC 3000  
QY 3148 GAGCACCGTGGCGCTTCGCGCAACCTCTGCGCTGGATGATCGTGAATTTCTTACGCG 3207  
Db 3001 GAGCACCGTGGCGCTTCGCGCAACCTCTGCGCTGGATGATCGTGAATTTCTTACGCG 3060  
QY 3208 CTGGTCAAGGCGCGAGACTTTGATTCGCGCTGCGAGATGTGGCACCCCACTGCTGTT 3267  
Db 3061 CTGGTCAAGGCGCGAGACTTTGATTCGCGCTGCGAGATGTGGCACCCCACTGCTGTT 3120  
QY 3268 CGCTCGATGCGATCTCTGAGCAGACGATAGGGTATGCGCAATGTTGTGCGCAACGTC 3327  
Db 3121 CGCTCGATGCGATCTCTGAGCAGACGATAGGGTATGCGCAATGTTGTGCGCAACGTC 3180  
QY 3328 AAGCGCCAGATCGGCCAATGCGTGTGCGTGAACCGCTCCGTTGAGTCTGTCAACCGAAC 3387  
Db 3181 AAGCGCCAGATCGGCCAATGCGTGTGCGTGAACCGCTCCGTTGAGTCTGTCAACCGAAC 3240

3388 GCAGAAAGGAGAGATTCCTCAACAAAGGCGCATGTTGCTGCACCATTCGCTGGTGTGTC 3447  
3241 GCAGAAAGGAGAGATTCCTCAACAAAGGCGCATGTTGCTGCACCATTCGCTGGTGTGTC 3300  
3448 ACCGTGACTGTTGCTGAAGTGATGAGGTCAAGCTGAGATGCGATCGCAATCATCGAG 3507  
3301 ACTGTGACTGTTGCTGAAGTGATGAGGTCAAGCTGAGATGCGATCGCAATCATCGAG 3360  
3508 GCTATGAAGATGAAGCAACAATCACTGCTTCGTTGACGGCAAAATCGATCGCTGTGG 3567  
3361 GCTATGAAGATGAAGCAACAATCACTGCTTCGTTGACGGCAAGATGGAAGCGTGTGG 3420  
3568 GTTCCTGCTCCACGAAGTGAAGTGGCGACTTGATCGTCTGCTTCCTAA 3621  
3421 GTTCCTGCTCCACGAAGTGAAGTGGCGACTTGATCGTCTGCTTCCTAA 3474

RESULT 14  
AP005216 308750 bp DNA linear BCT 10-JUL-2003  
DEFINITION Corynebacterium efficiens YS-314 DNA, complete genome, section 3/11.  
ACCESSION AP005216 BAO00035  
VERSION AP005216.1 GI:23492422  
KEYWORDS Corynebacterium efficiens YS-314  
SOURCE Corynebacterium efficiens YS-314  
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1  
AUTHORS Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and Gojobori, T.  
TITLE Comparative Complete Genome Sequence Analysis of the Amino Acid Replacements Responsible for the Thermostability of Corynebacterium efficiens  
JOURNAL Genome Res. 13 (7), 1572-1579 (2003)  
MEDLINE 22723752  
PUBMED 12840036

REFERENCE 2 (bases 1 to 308750)  
AUTHORS Kawarabayasi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and Director-General of Biotechnology Center.  
TITLE Direct Submission  
JOURNAL Submitted (17-MAY-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan (E-mail: bioelite.go.jp, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)

COMMENT Kawarabayasi, Y. is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, 305-8566 Japan  
Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan  
Ito, T. is at the Japan Biological Information Research Center, Koto-ku, Tokyo, 135-0064 Japan  
Yamagishi, A. is at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392 Japan  
Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co., Inc., Kawasaki, Kanagawa, 210-8681 Japan  
The other authors are at the National Institute of Technology and Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.  
Location/Qualifiers  
1. 308750  
/organism="Corynebacterium efficiens YS-314"  
/mol\_type="genomic DNA"  
/strain="YS-314"  
/db\_xref="taxon:196164"  
155. 598  
/notes="CE0586, similar to AE007159-8|AAK47889.1| percent identity: 63 in 145 aa"  
/codon\_start=1  
/transl\_table=11  
/product="putative 50S ribosomal protein L13"  
/protein\_id="BAC17396.1"

FEATURES  
source  
1. 308750  
/organism="Corynebacterium efficiens YS-314"  
/mol\_type="genomic DNA"  
/strain="YS-314"  
/db\_xref="taxon:196164"  
155. 598  
/notes="CE0586, similar to AE007159-8|AAK47889.1| percent identity: 63 in 145 aa"  
/codon\_start=1  
/transl\_table=11  
/product="putative 50S ribosomal protein L13"  
/protein\_id="BAC17396.1"

/db\_xref="GI:23492423"  
/translation="MSTYHPKSGDITREKVVVIDATDVVLGRLATHAADLLRGKGPILF  
APVDCGDVLIINADKVAITSNKRREMYRHSGYPGGLKSMTLGRSLLEHFERVLE  
ESTIGPMHNNKLTAAASKVLHVFAGSEHPFAAQKPEYIEIKVQA"  
598. 1146  
/notes="CE0587, similar to AL583918-48|CAC29873.1| percent identity: 59 in 159 aa"  
/codon\_start=1  
/transl\_table=11  
/product="putative 30S ribosomal protein S9"  
/protein\_id="BAC17397.1"  
/db\_xref="GI:23492424"  
/translation="MSEPIQENVESNVDAADIAATAATAEFTTNGDAISTSTEE  
ETFEAPAVLDGPIWTVGRKRAIVRVVAGSGEITCNGRTLEDVFPNKLHQLLKA  
PLVLGREGQFDIHANLGGGPTGQAGAFRLAIALARNAYNPAERETLTKAGPLTRDA  
RAVERKAGLHKARRAPQYSKR"  
1282. 2670  
/notes="CE0588, similar to AL031317-14|CAA20392.1| percent identity: 62 in 448 aa"  
/codon\_start=1  
/transl\_table=11  
/product="putative phospho-sugar mutase"  
/protein\_id="BAC17398.1"  
/db\_xref="GI:23492425"  
/translation="MCRAGGRWTCIMLGWTRLFGTDGVRGLANKTLTAPLALKGA  
AAHLVAGTRPHGRERPAIVGRDPRVSGEMLAALAGMASRGVDLVRGVITGPVA  
FLTDDGADMGVMSHSPNDNGIKFFSAGHKLPDEVEDIEIRWMDLFEETGPG  
HGIRVEEAPDARGYLQHLADAPDLSGTTVVDAANGAAVIAPOAYEAGAKV  
IAIHTPNAYINKEGCTHMDQIAAVLEHGDGLGADHGDADRLCLVDGNIYDG  
DQMAILAIAMKENSELRKNTLVATVMSNLGLKAMEKADIQRTTKVGDVRYLLELN  
ACGPAALGGEQSGHIVLPDGTGTGTLTGLSLMARMAATKPLSELAAAMTVLPQVLI  
NVFVADKSSIMKSANQAAVAABEELGGTGRVLRPSGTEELFRVWVAAAEQEQR  
VAGLAAVVAEA"  
2742. 3071  
/notes="CE0589"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17399.1"  
/db\_xref="GI:23492426"  
/translation="MLRGIVMSGLNVEGALTAAETMLGEATHQORHRVAPBP  
SNAGQGSFVGAQVATMBELHVRKLDVDAVEATAAAREVGVFADADAGIRGLG  
QVDSGEGR"  
3068. 4462  
/notes="CE0590"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17400.1"  
/db\_xref="GI:23492427"  
/translation="MSASISSYGSAYSQQAASRGTYPGPSIPERVLLKAYETITGLD  
TDQLRSGRGAAGDSTGILLGLKPLGRMGVAGGMIINELVRSWVDNRNEAE  
EVAABRAADAIDTTRISDDGVBAIILGOLAGIIAIESNHLSTLDPVEHPAFIATV  
QAGAEIINDAAAMVLGCGDRDRAIRSCYCALIEHGRVCEAPAGVAPVAGGGGGG  
AGTIPSPVTPPTTSGATPGTAPASDPPTADSPAGTATGTEADCDPEPGDEK  
PVSPAPPEDEDTVEPSPEDDDWEHLITGASIGIGLLVGVGVALVDFELGVAQ  
TMGDDAQMPPVPSGAPPLATPTVEVPFAPAPESAPFPNPTFPKTFPASLHLVA  
EPPPKPGVMVQAAAATAAQVQSEPTVGAPVQVQYQQAATPAAPVAPPTAAAPAGNSGDA  
AGDAPGRVRRVGGW"  
4456. 4716  
/notes="CE0591"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC17401.1"  
/db\_xref="GI:23492428"  
/translation="MVAVEYEVFAQFRQRTAARMAEFKALAKAQGEVAKTSFGFPG  
RQAAMSPPOVTPASVPARTPOQGRGRGRGPVQSILRRA"  
complement(4743. 5663)  
/notes="CE0592, similar to AL583918-51|CAC29878.1| percent identity: 44 in 254 aa"  
/codon\_start=1  
/transl\_table=11

```
/product="conserved hypothetical protein"
/db_xref="GI:23492429"
/translation="MAHPGRRIILPSGVTLSTILDRVAENLKHLKSLKSGRPHRVVVG
DLVAGLPKITYPAEGNGVGVPAFGHDMKPIKHVHOTLRHLASWGIAVAAPANTKG
FLPHRGPADLSDALQILAGLVKAGKVTNPGGLGVGHGMMGVAVLAAANRKG
RVALVPASTSPSCVDAAAPAKVAGLIICTADLGMEDAGDPAKVAARWGVVVEL
DNGNQFSEDTWFKLVGVGRPQTACQELARGLLIGFLHLQLAGRSYKAFSDPAE
AKVTISWGIDLEKAPQDESPLFFLSN"
gene 5737.6840
CDS /gene="alr"
/gene="alr"
/EC number="5.1.1.1"
/note="CE0593, similar to AX063847-1|CAC25165.1| percent
identity: 68 in 362 aa"
/codon_start=1
/transl_table=11
/product="alanine racemase"
/protein_id="BAC17403.1"
/db_xref="GI:23492430"
/translation="MNLITRIIDLDAHTNRLKNRVGAALKMAVVKADVYNGMDR
VAPYANGADAFVATIAENALRATGITPILCNWSPEDQWAAAEVAGIDLAVIS
PRHARLVADPTSRALRVSKVDTALHRSVDEQDWDVFTLRLDCHGLEVGLFTH
LSCADEPNPTDHOAETFRRAIDRALGLVPEHLCNSPATLFRPDLHMDVVRPG
VALYGLPFIPLDHLRPAFTWAGAITVVKPIRKEGTSGYLTWRABDGLFVAVPAG
YADGVPRAAQDLHLYVHNGHDYPOQGVCMDFVIFLGDNPHTVAGDEAVIFGGTGM
SATELADALINVEYICRPTGRTVRDYGEGK"
6837.7340
CDS /note="CE0594, similar to AL583918-56|CAC29885.1| percent
identity: 50 in 156 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17404.1"
/db_xref="GI:23492431"
/translation="WKDPFASGTERLDTAAQTQAFGEELGRALBAGDVVVLDPGLGA
KKTFTQGLRGMQVRGVSTPFIARERSEVGGFTLIHLDAYRLGLBDDPDVSD
FIGALDSLDLTDLTAVVVAEWGGGLVEQITESYLLISIDRETAVENTDEPESARIFT
WQWRSRH"
7540.9132
CDS /note="CE0595, similar to AX066613-1|CAC26534.1| percent
identity: 41 in 530 aa"
/codon_start=1
/transl_table=11
/product="putative transport protein"
/protein_id="BAC17405.1"
/db_xref="GI:23492432"
/translation="WLDPLAANPLIALAVILAVGLAIGRISLFGVSLGAAVLIVLV
VSTLNPDIQIPAFVPLGLAMFVVYVIGISAGPAFFRPSRSGWKLTLFMTLLVSLTA
LAWLIRAFGIGAGAGAGMPAGSLSTPGWAAVVALMDPAQGDPPVIGSLAYEGAVL
GSILVAVGAKLKNHTEDAREGLVTEPLWGVTRIGISGISTIGDLPRLSQQII
ATRIVEDPHERLADPTLPIKPMELVINGTVNAVDAIRAIALGEGECDTKIEDTELVS
RFTVSNPDIVORTVAEDFVANGFMIRIQGDDEIIVPHRDVTNLNSDRVRVVAAPR
MGEVRFPLGDSKALGDVNLPLPAIGSLGLLGLAI FVPLPGDTTWYLGFGGPFVAG
LILGALNRGTGITWLPNHRNTISTIGLALFLAGVTSAGAGFRQALTDPQSFVYMG
VGFAITVTSALVCAVGMWLKIKWDESMVGMAAGATTNPALISYINDQTGDLANRY
ATVPTAMIGKILACQVLELL"
9183.10529
CDS /note="CE0596, similar to AX065751-1|CAC26115.1| percent
identity: 79 in 430 aa"
/codon_start=1
/transl_table=11
/product="putative prolyl aminopeptidase"
/protein_id="BAC17406.1"
/db_xref="GI:23492433"
/translation="WAPRGGTGWPDWGLTGGGVPVSTRGNLDGNTLSLTHLGLRL
DELITVPLTIDVADERTIDIFARIATPGGETLPIVFIQGGFNGNAPRSLPPLNP
SWLKAALHRYVWMLDQGTGLTSLVPSDRILEEHTAEIVELSHLEADGIVRDEAV
REYLGVRQWYVLGSGFGFTLHLYSTLHPSLDLDVFTTGLSLDAIRPAEDVYANCYR
MRYNSEQYRPPEDRVIGIQLVERARAGEIILPTEGVSESLRSLRSLHGLSNDGWL
DLHLLEDPRSNADLADLLPFGARNPLVYVHSSYADGVVTDWAAQRYVPEAF
REDPTLLTGEHVFPWADTVPALRPWKDVALALAAQEWPKLYDAALETSGAAGAV
```

	Query Match	63.0%	Score 2282.8	DB 1	Length 308750	
	Best Local Similarity	77.7%	Pred. No. 0			
	Matches 2810	Conservative	0	Mismatches 797	Indels 9	Gaps 4
QY	7	GGGGTTAGATCTCTGGGGGGTTTATTTCATTACCTTTTGGCTTGAAGTCTGTCAGGTTCAGGG	66			
DB	128583	GGGGTCTCGGGGGGGGGGCAATCTTTTTCACGGCAAGGTGTGTGAATTCCTCGGCACTC	128642			
QY	67	GAGTGTTCGGCAAAACATTTGAGAGGAAAAAACAACCATGATTTTGTGTTGGGGGAATCGG	126			
DB	128643	CCCGGCCCGCGGTAGAGAACGGAGCGAAACGGAAGCAATACGTGGTTCCTCGGACTGG	128702			
QY	127	GGGTTCACATATCTAGACGCGAGTCTGCTATCACCCCTT--GGCGGTCTCTTGTGTAAAG	184			
DB	128703	CGGTTCAGATGTTCTGAAG-AGTGNCTGCCATCACCAACAGGTGTGCTCTCGTCAAG	128761			
QY	185	GAATAATTACTCTAGTGTGCACTACACATCTTCAACGCTTCAGCATTTCAAAAGATCT	244			
DB	128762	GAACAAAAAATGTGGTTTACAACA---ACACCTCTCCAGCTCCGCGCTTCAAAAGATCC	128818			
QY	245	TGCTAGCAAAACCGCGCGCAAAATCGCGGTCCGTGCTTTCGCTGCAGCACTCGAAACCGGTG	304			
DB	128819	TGCTGGCCAAACCGAGGTGAATTCGCGGTGCGGCAATTCGCGCGCCTACGAGACCGGG	128878			
QY	305	CAGCCACCGGTAGCTATTTACCCCGGTGAAGATCGGGGATCAATTCACCGCTCTTTTGCTT	364			
DB	128879	CCGCAACCGTGGCCATCTACCCCGGAGGACCGTGGCTCTCTCCACCGCTCTCTCGCT	128938			
QY	365	CTGAGCTGTTCGCATTTGGTACCGAAGGCTCACAGTCAAGCGTACCTGCAGCATCGATG	424			
DB	128939	CCGAGCGGTGAGGATCGAAACCGAGGCTTACCCGTCAAGCGTACCTCGATTTGATG	128998			
QY	425	AAATATTCGCTCAGCTTAAAGTAAAGCAGATGCCATTTACCCGGATACCGCTTCC	484			
DB	128999	AGATCATCAACCGCCCAAGAGGTGAAGCGACCGGTCTACCCGGGTATGCTTTCC	129058			
QY	485	TGCTGAAAATCCAGCTTTCGCCGAGGTGTGCGGAAACGGCAATTAATTTATTTGCC	544			
DB	129059	TTTCGAAAATCCAGCTTCGCGCTGCGGTGAATGCGCGGAGACGGCAATTAATTTATTTGCC	129118			
QY	545	CAACCCCGAGGTCTTCTGATCTCACCGGTGATAAGTCTCGCGCGGTAAACCGCGCGAAGA	604			
DB	129119	CCACCCCGAGGTCTCGACCTCACGGGCGACAAGTCTCGCGCGCGCGAAGA	129178			
QY	605	AGCTGTCTGCGAGTTTGGCGGAAATCCACCCCGAGCAAAACATCATGATGATCGTTA	664			
DB	129179	AGCGCGGCTGCGGTGCTGCGGGAATCCACCCCGAGCAACCATCATGATGATCGTTA	129238			
QY	665	AAAGCGCTCAAGCCAGACTTACCCCATCTTTGTGAAGCAGTTCGCGGTGTGCGCGAC	724			
DB	129239	AGAGTCCGAGGGCGAGCACTTACCCGATCTTCTCAGGCGCTCGCAGGTGTGCGCGGC	129298			
QY	725	GGGTTATGCTTTTGTGTTTCACTGATAGCTTCGCAAAATAGCAACAGAGATCTC	784			
DB	129299	GTGTTATGCGGTTCGTCGAGAAAGCCGAGGACCTCGGTGAGTGTGCGGAGGCTTCC	129358			
QY	785	GTGAAGCTCAAGCGGCTTTTCGCGCATGCGCGGTATATGTCGAAAGTGTGTGATTAAAC	844			
DB	129359	CGAGGCGAGGCGCGCTTTTCGCTGAGCATTCGCTCTACGTCGAAAGGCGGTGATCAAC	129418			
QY	845	CTCAGCATATTAAGTGCAGATCTTTGGCGATCACTTGGAGAAATGTGTACACCTTTATG	904			
DB	129419	CCGAGCACAATCGAGGTGCAGATCTCTCGGTGATCACACCGCGCATGTCTATCCACCTGTATG	129478			
QY	905	AAAGTCACTGCTCACTGCAGCGCTGTCAACAAAAAGTTGTGCAAAATTTGGCCAGCACAGC	964			
DB	129479	AACCGACATGTTCCCTGACGCGCCGCCACAGAGGTGTGTGAGATCGCACCTGTGCCAGC	129538			
QY	965	ATTTCGATCAGAACTGCGTGTATGCAATTTGTCCGATGATGCAATTAAGTTCTCCCGCTCA	1024			
DB	129539	ACCTCGACCCGAGCTGCGCGACCGCATCTGTGCGCATGCCGTGAAGTTCTGCAAAATCCA	129598			

QY 1025 TTGGTTACAGGGCGGGGAAACCGTGGAAATCTTGGTGCATGAAAGGGCAACCAAGTCT 1084  
Db 129599 TCGGATACAGGGCGGGGCAACCGTGGAAATCTTGGTGCATGAAAGGGGCAACCAAGTCT 129658  
QY 1085 TCATCGAATGAACCAACGATATCCAGGTGGAGACACCGTGGAGTGAAGAAATGACCGAGG 1144  
Db 129659 TCATGGAGATGAACCCCGCATCCAGGTGGAGACACCGTGGAGTGAAGAAATGACCGAGG 129718  
QY 1145 TGACCTGTGTGAAGGGCGAGATGCGGTTGGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1204  
Db 129719 TGACCTGTGTGAAGGGCGAGATGCGGTTGGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 129778  
QY 1205 TGACCCCAAGATGAAGATCAAGACCCAGGTGGAGACACCGTGGAGTGAAGAAATGACCGAGG 1264  
Db 129779 TGACCCCAAGATGAAGATCAAGACCCAGGTGGAGACACCGTGGAGTGAAGAAATGACCGAGG 129838  
QY 1265 ATCCAAACAAAGGCTTCGCGCCAGATACCGGAACTATACCGCGTACCGCTACCGAGCG 1324  
Db 129839 ATCCAAACAAAGGCTTCGCGCCAGATACCGGAACTATACCGCGTACCGCTACCGAGCG 129898  
QY 1325 GAGCTGCGGTGTGTCTTGACGGTGGAGTCAAGTGGTGGGGAATACCGGAACTATACCGGAA 1384  
Db 129899 GAGCTGCGGTGTGTCTTGACGGTGGAGTCAAGTGGTGGGGAATACCGGAACTATACCGGAA 129958  
QY 1385 ACTCCATGCTGGTGAAGATGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1444  
Db 129959 ACTCCATGCTGGTGAAGATGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 130018  
QY 1445 CACAGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1504  
Db 130019 CACAGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 130078  
QY 1505 GTGCGTGTGTGGCGGAGAGGATTCACCTTCAAGCGCATCCGACCGGATTCATTTGCGG 1564  
Db 130079 GTGCGTGTGTGGCGGAGAGGATTCACCTTCAAGCGCATCCGACCGGATTCATTTGCGG 130138  
QY 1565 ATACCGCGACCTTCCTTCAGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 1624  
Db 130139 ATACCGCGACCTTCCTTCAGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 130198  
QY 1625 ACTTGGCAGATGTCACCGGTGAACCAAGCTCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1684  
Db 130199 ACTTGGCAGATGTCACCGGTGAACCAAGCTCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 130255  
QY 1685 CTATCATAGCTTCCTTCAGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 1744  
Db 130256 CTATCATAGCTTCCTTCAGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 130315  
QY 1745 TGAAGCAGCTTGGCGGCGGCGGTTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1804  
Db 130316 TGAAGCAGCTTGGCGGCGGCGGTTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 130375  
QY 1805 TTACTGATACCACTTTCGCGGATGACACCGAGTTCCTTGGTGGTGGTGGTGGTGGTGGTGGTGG 1864  
Db 130376 TTACTGATACCACTTTCGCGGATGACACCGAGTTCCTTGGTGGTGGTGGTGGTGGTGGTGGTGG 130435  
QY 1865 TCGCAGCTGAAGCTGGCGGAGAGCGGTCGCAAGCTGACTCTGAGCTTTTGTCCGTTGG 1924  
Db 130436 TCGCAGCTGAAGCTGGCGGAGAGCGGTCGCAAGCTGACTCTGAGCTTTTGTCCGTTGG 130495  
QY 1925 AGGCTGGGCGGCGGAGACCTAAGATGGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1984  
Db 130496 AGGCTGGGCGGCGGAGACCTAAGATGGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 130555  
QY 1985 ACAGGCTGACGAGTGGCGGAGCGGATGCGGAATGTAAGATTCAGATGCTGCTTCGCG 2044  
Db 130556 ACAGGCTGACGAGTGGCGGAGCGGATGCGGAATGTAAGATTCAGATGCTGCTTCGCG 130615  
QY 2045 GCGCGAACAACCGTGGAGATACCCCGGTAACCGAGTTCGCTGGTGGTGGTGGTGGTGGTGGTGG 2104  
Db 130616 GCGCGAACAACCGTGGAGATACCCCGGTAACCGAGTTCGCTGGTGGTGGTGGTGGTGGTGGTGG 130675  
QY 2105 AAGCTGCCAGTCCGGGCGGAGACATCTTCGGCATCTTCGACGCGCTTAACGAGCTCTCC 2164

Db 130676 AGCCGCCAAGTCCGGTGGAGCATCTTCGGCATCTTCGACCGCTCAACGACATCTCC 130735  
QY 2165 AGATGGCTCCAGCAATPCGAGCGAGTCTTGAGACCAACCAACCGCGGTAGCCGAGGTGGTGA 2224  
Db 130736 AGATGGCTCCAGCAATPCGAGCGAGTCTTGAGACCAACCAACCGCGGTAGCCGAGGTGGTGA 130795  
QY 2225 TGCGTTATTTCTGGTGTCTCTGTATCCAAATGAAGAGTCTTACACCTGGATTAATCTAC 2284  
Db 130796 TGCGTTATTTCTGGTGTCTCTGTATCCAAATGAAGAGTCTTACACCTGGATTAATCTAC 130855  
QY 2285 TAAAGATGGCAGAGGAGATCGTCAAGTCTTGGCGCTCACAATCTTGGCCATTAAGGATATGG 2344  
Db 130856 TAAAGATGGCAGAGGAGATCGTCAAGTCTTGGCGCTCACAATCTTGGCCATTAAGGATATGG 130915  
QY 2345 CTGGTCTGCTTGGCCAGCTGGGTAAACCAAGTGGTCAACCGCATGGCGCGGTGAATTCG 2404  
Db 130916 CTGGTCTGCTTGGCCAGCTGGGTAAACCAAGTGGTCAACCGCATGGCGCGGTGAATTCG 130975  
QY 2405 ATCTGCCAGTGCAGTGCACACCCAGCACATCGCGGTGGCCAGTGGCAACTACTTTG 2464  
Db 130976 ATCTGCCAGTGCAGTGCACACCCAGCACATCGCGGTGGCCAGTGGCAACTACTTTG 131035  
QY 2465 CTGAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGTTCGCGACCACTGTCTGGCACCA 2524  
Db 131036 CTGAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGTTCGCGACCACTGTCTGGCACCA 131095  
QY 2525 CTTCGAGCCTATCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2584  
Db 131096 CTTCGAGCCTATCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 131155  
QY 2585 GTTGGAGCTTCGAGGCTGTTTCTGACCTCGAGCGTACTTGGAGAGAGTGGCGGAGTGT 2644  
Db 131156 GTTGGAGCTTCGAGGCTGTTTCTGACCTCGAGCGTACTTGGAGAGAGTGGCGGAGTGT 131215  
QY 2645 ACCTGCCATTTGAGTCTGGAACCCAGGCGCCAAACCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 2704  
Db 131216 ACCTGCCATTTGAGTCTGGAACCCAGGCGCCAAACCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 131275  
QY 2705 CAGCGGACAGTGTGTCACCTGCTGTCACAGCGCCACCGCATCTGGGCTTGGCGGATTCGTT 2764  
Db 131276 CAGCGGACAGTGTGTCACCTGCTGTCACAGCGCCACCGCATCTGGGCTTGGCGGATTCGTT 131335  
QY 2765 TCGAAGCTATCGAAGCAACTACGAGCGCTTAATGAGATGCTGGGACGCGCCAAACCAAG 2824  
Db 131336 TCGAAGCTATCGAAGCAACTACGAGCGCTTAATGAGATGCTGGGACGCGCCAAACCAAG 131395  
QY 2825 TCACCCCATCTCCAGGTTTGGGAGCTTCGCACTTCCACCTCGTTGGTGGCGGTTGG 2884  
Db 131396 TCACCCCATCTCCAGGTTTGGGAGCTTCGCACTTCCACCTCGTTGGTGGCGGTTGG 131455  
QY 2885 ATCCAGCAGACTTGTGCGGATCCCAAAAGTACGACATCCCGAGACTCTGTATCGCGT 2944  
Db 131456 ATCCAGCAGACTTGTGCGGATCCCAAAAGTACGACATCCCGAGACTCTGTATCGCGT 131515  
QY 2945 TCCTTCGCGCGGAGCTTGGTAAACCTTCAGTGGTGGCGAGAGCCATCGCGACCGCG 3004  
Db 131516 TCCTTCGCGCGGAGCTTGGTAAACCTTCAGTGGTGGCGAGAGCCATCGCGACCGCG 131575  
QY 3005 CACTTGAAGCGCGCTCCGAGGCAAGCAGCTCTGACGGAAGTTCCTGAGGAAGAGCAGG 3064  
Db 131576 CACTTGAAGCGCGCTCCGAGGCAAGCAGCTCTGACGGAAGTTCCTGAGGAAGAGCAGG 131635  
QY 3065 CGCACCTCGAGCTGTGATGATTCGAAGGAACGTGCAATAGCTCAACCGCTCTGTGTTC 3124  
Db 131636 CGCACCTCGAGCTGTGATGATTCGAAGGAACGTGCAATAGCTCAACCGCTCTGTGTTC 131695  
QY 3125 CGAAGCCAAACGAGAGTTCCTCGAGCAGCGTGGCGCTTGGCGCAACACCTCTGCGCTGG 3184  
Db 131696 CGAAGCCAAACGAGAGTTCCTCGAGCAGCGTGGCGCTTGGCGCAACACCTCTGCGCTGG 131755  
QY 3185 ATGATCGTGAATCTTCTACGGCTGCTCGAAGCGCGGAGACTTTGATCGCTGCGCAG 3244

Db	131756	ATGACCGCGAGTTCTTCTACGGCTTGTAAGAGGAGCGCTGAGGAGCTGATCCGACTGACCG	131815
Qy	3245	ATGTGCGCACCCCACTGTTCTTCCCTCTGGATGCGATCTCTGAGCCAGACGATAAGGGTA	3304
Db	131816	GTGTGTCCACCCCGATGTTGTCTGCTGCTGATGCGGTGTCGAACCGGATGACAAAGGCA	131875
Qy	3305	TGCGCAATGTGTGGCCACGTCACCGCCAGATCCGCCCAATGCGTGTGCGTGACCGCT	3364
Db	131876	TGCGCAACGTTGTGTCAACGTCACCGCCAGATCCGCCCGATCAAGGTGCGGACCGTT	131935
Qy	3365	CGGTTGAGTCTGTCAACCGCAACCGCAGAAAGGAGGATTCCTCAACAAAGGCGCATGTTG	3424
Db	131936	CGTGGAGTCCGTCAACCGCCACCGCGGAGAGCGCGATGCCACCAACAGGCGCATGTG	131995
Qy	3425	CTGCACCAATTCGTGTGTTCTACCGTGACTGTGTCTGAGGTGATGAGTCAAGGCTG	3484
Db	131996	CCGCACCAATTCGCGGTGTGTCACCGTGACCGTGCAGGAGGTGATGAGATCAAGGCTG	132055
Qy	3485	GAGATGCAATCGCAATCATCAGAGCTATGAAGATGGAAGCAACCAATCACTGCTTCTGTTG	3544
Db	132056	CGACGCGGTGCCCATCATGAGGCCATGAAGATGGAGGCCACCATCAACGCGCTGTG	132115
Qy	3545	ACGGCAAAATCGATCGCTTGTGTTCTGCTGTCGCAACGAAGGTGGAAGGTGGCGACTTGA	3604
Db	132116	ACGGTGTCTATGACCGCGCTGCTGTGTCGCGCCGCCACCAAGGTGCGAGGCGCGACCTCA	132175
Qy	3605	TCGTCGCTGTTTCTTA 3620	
Db	132176	TCGTGTGCTGTCCTTA 132191	
RESULT 15			
BD094153			
LOCUS	4013 bp	DNA	linear PAT 27-AUG-2002
DEFINITION	Genes for heat resistant enzymes of amino acid biosynthetic pathway derived from thermophilic coryneform bacteria.		
ACCESSION	BD094153		
VERSION	BD094153.1	GI:22639741	
KEYWORDS	WO 0125447-A/11.		
SOURCE	Corynebacterium thermoaminogenes		
ORGANISM	Corynebacterium thermoaminogenes		
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.		
AUTHORS	Hirano, S., Nonaka, G., Matsuzaki, Y., Akiyoshi, N., Nakamura, K., Kimura, E., Osumi, T., Matsui, K., Kawahara, Y., Kurahashi, O., Nakamatsu, T. and Sugimoto, S.		
TITLE	Genes for heat resistant enzymes of amino acid biosynthetic pathway derived from thermophilic coryneform bacteria		
JOURNAL	Patent: WO 0125447-A 11 12-APR-2001.		
	AJINOMOTO CO INC, SEIKO HIRANO, GEN NONAKA, YUMI MATSUZAKI, NAOKI AKIYOSHI, KANAE NAKAMURA, EIICHIRO KIMURA, TSUYOSHI OSUMI, KAZUHIKO MATSUI, YOSHIO KAWAHARA, OSAMU KURAHASHI, TSUYOSHI NAKAMATSU, SHINICHI SUGIMOTO		
COMMENT	OS Corynebacterium thermoaminogenes		
	PN WO 0125447-A/11		
	PD 12-APR-2001		
	PF 04-OCT-2000 WO 2000JP006913		
	PR 04-OCT-1999 JP 99P 282716, 01-NOV-1999 JP 99P 311147 PR		
	21-APR-2000 JP 00P 120687		
	PI SEIKO HIRANO, GEN NONAKA, YUMI MATSUZAKI, NAOKI AKIYOSHI, KANAE NAKAMURA,		
	PI EIICHIRO KIMURA, TSUYOSHI OSUMI, KAZUHIKO MATSUI, YOSHIO KAWAHARA,		
	PI OSAMU KURAHASHI, TSUYOSHI NAKAMATSU, SHINICHI SUGIMOTO PC		
	C12N15/60, C12N15/54, C12N15/53, C12N15/31, C12N15/56, C12N9/88, PC		
	C12N9/12, PC C12N9/04, C07K14/34, C12N9/26, C12N13/04		
CC			
FF	Key Location/Qualifiers		
FT	CD5 Location/Qualifiers (319) . . (3735).		
FEATURES	source		
	1. . 4013		

		/organism="Corynebacterium thermoaminogenes"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:161484"	
ORIGIN		Query Match	
		Best Local Similarity 63.0%; Score 2279.6; DB 6; Length 4013;	
		Matches 2808; Conservative 0; Mismatches 799; Indels 9; Gaps 4;	
Qy	7	GGGGTTAGATCCTCGGGGGTATTATTCATTCATCTTTGGCTTGAAGTCTGTCAGGTGACGG	66
Db	129	GGGGTCTCGGGCGGGGGCATTTCTTTTCCGCGCAGGTGGTGAATTCGCGAGTCACTC	188
Qy	67	GAGTGTTCGCGAAACATTTGAGAGGAGAAAACAAAACCGATGTTTGAATGGGGGAATCG	126
Db	189	CCGCGCGCGGTAGAGAACGAGGAGAAAACGGAAGCAATACGTGGTTCGCGACTGG	248
Qy	127	GGGTTACGATCTAGGACGCGAGTGTCTATCACCTT--GGCGGTCTCTTTGTTGAAG	184
Db	249	CCGTTACGATGTTCTGAAG-AGTGACTGCCATCACCAACAGGCTGGTCTCTCGTGAAG	307
Qy	185	GAATTAATCTAGTGTGCTCACATCTTCAACGCTTCCAGCATTCAAAAGATCT	244
Db	308	GAACAAAACCTGTGTTACAACA---ACACCTTCCACGCTGCGGCGTTCAAAAGATCC	364
Qy	245	TGTTAGCAAAACCGCGCGAAATTCGCGGTCCGTGCTTCCGTGACGACCTCGAAAACCGTG	304
Db	365	TGTTGGCCAAACCGAGGTGAATTCGCGGTGCGAGCATTCGCGCGCCCTACGAGACCGGG	424
Qy	305	CAGCCAGGTAGTATTTTACCCCCGTGAGATCGGGGATCATTTCCACCGCTCTTTTGTCT	364
Db	425	CCGCAACCGTGGCCATCTACCCCCCGGAGGACCGTGGCTCTTCCACCGCTCTTCTGCT	484
Qy	365	CTGAAGCTCTCCGCAATTTGGTACCGAAGGCTCACAGTCAAGCGCTACCTGGACATCGATG	424
Db	485	CCGAGCGGTGAGGATCGGAACCGAGGGCTCACCCGTCAAGCGCTACCTCGATATTGATG	544
Qy	425	AAATATCGGTGAGTATAAAGTTAAGCAGATGCCATTTACCGGGATACCGGTCTCC	484
Db	545	AGATCATCAACCCCGCAAGAGGTGAAGCGGACCGCGTCTACCCGGGTATGTTTCTCC	604
Qy	485	TGTTGAAAATCCCAAGCTTGCCCGCGAGTGTGCGGAAAACCGGCATTACTTTATTGGCC	544
Db	605	TTTCGAAAATGCCAGCTCGCGCGTGAATGCGCGAGAACGCGATTACTTTTCATCGGTC	664
Qy	545	CAACCCCGAGGTTCTTGATCTACCGGTGATAAGTCTCGCGGTAAACCGCGGGAAGA	604
Db	665	CCACCCCGAGGTGCTCGACCTCACGGGCGACAACTCCAAAGGCTGTGTCGCGGGAAGA	724
Qy	605	AGGCTGCTGCGCAGTTTGGCGGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTA	664
Db	725	AGGCGGGTTCGGTGTCTGGCGGATCCACCCCGAGCACCGACATCGATGAGATCGTCA	784
Qy	665	AAAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAGCAGTTCGCGGTGTTGGCGGAC	724
Db	785	AGAGTCCGAGGGCGAGACCTACCCGATCTTCGTCAAGCGCTCGCAGGTGTGCGGGGC	844
Qy	725	GGGTTATGCTTTTGTGTTGCTTCACTGTAGCTTCGCAAAATTAGCAACAGAGCATCTC	784
Db	845	GTGGTATGCGGTTCGTGAGAAAGCCGAGACCTGCGTGAAGTTCGCGAGGAGGCTCC	904
Qy	785	GTGAAGCTGAAGCGCTTTTCGCGGATGGCGCGGTATATGTGAAACGTGTGTGATTAAAC	844
Db	905	GCAGGCGGAGGCGCTTTTCGGTGAAGCTTCGTTACGTCGAAACGGGCGGTGATCAAC	964
Qy	845	CTCAGCATATTAAGTGCAGATCTTGGCGATCACACTGGAGAGTTGTACACCTTTATG	904
Db	965	CCGACCATTCGAGGTGCAATCTTCGGTGAATCACCGGCGATGTCTATCCCTGTATG	1024
Qy	905	AACGTACTCTCACTGACGCTGTTCACCAAAAAGTTGTGAAATTTGGCCAGACACG	964
Db	1025	AACGGACTTTCCTGACGCGCGCCACAGAAAGTTCGTGGAGATCGCACCTGCCAGC	1084



965 ATTTGATCCAGAACTGGCTGATCGCATTTTGTGCGGATGCGAGTAAAGTTTTCGCGCTCCA 1024  
1085 ACCTGACCCGAGCTGCGGACCGCATCTGTGCGGATGCGGATGCTTGCATATCCA 1144  
1025 TTGTTTACCAAGCGCGGAAACCGTGGAAATTTCTGTGATGAAAGGCGCAACACAGTCT 1084  
1145 TCGGATACCAAGCGCGGCAACCGTGGAAATTTCTGTGATGAAAGGCGCAACACAGTCT 1204  
1085 TCATCGAAATGAACCAACCGATTCAGGTTGAGACACCGTGTGATGAAAGTTCACCGAGG 1144  
1205 TCATTCAGATGAACCCCGCATCCAGGTGGAACACACCGTGTGACCGGAGGTTCACTCCG 1264  
1145 TGACCTGTGTGAAGCGGAGATGCGCTTGGCTGTGCTGCTGCAACCTTGAAGGAATTTGGTTC 1204  
1265 TCGACCTGTGTGAAGCGGAGATGCGCTTGGCTGTGCTGCTGCAACCTTGAAGGAATTTGGTTC 1324  
1205 TGACCCAGATGAAGTCAAGAACCGGTCAGCACTGCGGTCGCGATACCAACGGAAG 1264  
1325 TGACCCAGGACAGATCAACCAACCGTGTGCGGCTGCGTGTGATGCGGTCATCAACCGAGG 1384  
1265 ATCCAAACAGGCTTCCCGCCAGATACCGGAACATACCGGTCACCGCTCACCGAGCG 1324  
1385 ACCGCTCAACAACTTCCGCGCGACACCGGTGTGATGCGGTCACCGCTCACCGGCTG 1444  
1325 GAGCTGGCTGTGTGACGCTGACGCTGAGCTGAGCTGCGGTGCGGAATACCGGACACTTTG 1384  
1445 GTGCGGTGTGCTGTGACGCGGAGCGGCTGCGGCGGAGATCAACCGGACACTTTGCG 1504  
1385 ACTCCATGCTGTGAAATGACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1444  
1505 ATTCATGCTGTGCAAGATGACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1564  
1445 CACAGCGCGGTTGGCTGAGTTTCAACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1504  
1565 CCCAGCGCGCTTGGCGGAGTTCAACGCTTCCGCGTGGCGGACCAACATCGGCTTCTGTC 1624  
1505 GTGCTGTGCTGCGGAGGAGCTTCACTTCCAGCGCATCCGACCGGATTCATGCTGCTG 1564  
1625 GTGCTGTGCTGCGGAGGAGCTTCAACAGGCGGATGCGGCTGCTGCTGCTGCTGCTGCTG 1684  
1565 ATCAACCGGACCTTCTTCAAGGCTTCAACCTGCTGATGATGAGCAGGAGCGCATCTGCTGAT 1624  
1685 CCCACGACGCTTCTTCAAGGCTTCAACCTGCTGATGATGAGCAGGAGCGCATCTGCTGAT 1744  
1625 ACTTGGCAGATGTCAGGTCAGGAGCTTCACTTCCAGCGCATCCGACCGGATTCATGCTG 1684  
1745 ACTTGGCGGATGTCAGGTCAGGAGCTTCAACAGGCGGATGCGGCTGCTGCTGCTGCTGCTG 1801  
1685 CTATCGATAGCTGCTTCAACATCAAGGATGTCGCACTGCGGATGCGGCTTCCGCTGACCGG 1744  
1802 CGATAGAGAGCTGCGGAGTGGAGAACATCCGCTGCGGATGCGGCTTCCGCGACCGGCT 1861  
1745 TGAAGCAGCTTGGCCGACCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1804  
1862 TGAAGCAGCTTGGCCGAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1921  
1805 TTACTGATACCACTTCCGCGATGTCACAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1864  
1922 TCACCGACACCACTTCCGCGATGTCGCAACGATCCCTTCTGCGGACCCGCGTGGCTCT 1981  
1865 TCGCATGGAAGCTGCGGCGAGGCGGCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1924  
1982 TCGCGCTGACCCCGGCGGCGGCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2041  
1925 AGGCTTGGGCGGCGGCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1984  
2042 AGGCTTGGGCGGCTGCAACCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2101  
1985 ACAGGCTGCAAGGCTGCGGAGGCGGCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2044  
2102 CAGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2161  
2045 GCCGCAACCGTGGGATACACCGGTCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2104

2162 GTCCGAACACGTCGCGGTACACCCGATACCCGATTCGCTGCTGCTGCTGCTGCTGCTGCTG 2221  
2105 AAGCTGCCAGCTTCCGCGTGGACATCTTCCGATCTTCCGCGCTTAAACGAGCTCTCC 2164  
2222 AGCGCGCAAGTCCGCTGCTGCACTCTTCCGATCTTCCGCGCTTAAACGAGCTCTCC 2281  
2165 AGATGGTCCAGCAATCGACGAGTCTCTGAGACCAACACCGGCTAGCCGAGTGGCTA 2224  
2282 AGATGGCGCGGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCG 2341  
2225 TGGCTTATTTCTGCTGATCTCTCTGATCCAAATGAAAGCTCTACACCTTGGATTTACTAC 2284  
2342 TGGCTTATTTCTGCTGATCTCTCTGATCCAAATGAAAGCTCTACACCTTGGATTTACTAC 2401  
2285 TAAAGTGCAGAGGAGATCTGCACTCTGCGCTCACAATCTTGGCCATTAAGGATATG 2344  
2402 TGAACCTGCGGAGAGATCTGCACTCTGCGCTCACAATCTTGGCCATTAAGGATATG 2461  
2345 CTGCTCTGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTT 2404  
2462 CCGGCTGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTT 2521  
2405 ATCTGCGAGTGCAGCTGCAACCCAGACACCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 2464  
2522 ACTGCGCGTGCATGTCACACCCAGACACCGCGCGCTGAGTGGGCACTTACTGCG 2581  
2465 CTGCGCTCAAGCTGCTGAGATGCTGTTGAGCGCTTCTCCGACCACTGCTCTGGCACA 2524  
2582 CCGCGCAACCGCGGCGGCTGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 2641  
2525 CTTCCGAGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2584  
2642 CTTCCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2701  
2585 GTTTGAGCTTCCGCGCTTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2644  
2702 GCTTCAACTGCGAGGCTGCTTCCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 2761  
2645 ACTGCGCTTGGAGTTCGGAACCCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 2704  
2762 ACTGCGCTTGGAGTTCGGAACCCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 2821  
2705 CAGCGGACAGCTTGTCCAACTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 2764  
2822 CCGCGCTGAGCTGCTTCCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 2881  
2765 TCGAATCTGATCGAAGCACTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2824  
2882 TCGAGCTCATCGAGGCTACTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2941  
2825 TCACCCCATCTCCAGGCTTGTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCG 2884  
2942 TCACCCCATCTCCAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3001  
2885 ATCCAGCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2944  
3002 GCGCGGAGGATTCGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCG 3061  
2945 TCGTCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3004  
3062 TCGTCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3121  
3005 CACTGGAAGCGCTGCGAAGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3064  
3122 CACTGGAAGCGCTGCGAAGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3181  
3065 CGCAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3124  
3182 CCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3241  
3125 CGAAGCAACCGAAGGCTTCTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3184



Db	3242	CGAAGCCGACGAGGAGTTCTTGAGCACCGTCGCGCTTCGGCAACACCTCCGCCCTGG	3301
Qy	3185	ATGATCGTGAATTTCTTACGGCTGTGCGAAGCCGCGAGACTTTGATCGCTGCCAG	3244
Db	3302	ATGACCGCGAGTTCTTCTACCGCTTGAAGGAGGACGTGAGGAGCTGATCCGACTGACCG	3361
Qy	3245	ATGTCCGACCCCACTGTTGTTCGCTTGGATGCGATCTCTGAGCCAGACGATAAGGTA	3304
Db	3362	GTGTGTCCACCCGATGGTGTTCGCTTGGATGCGGTGTCGGAACCGGATGACAAAGCA	3421
Qy	3305	TGCGCAATGTTGTGGCCCAACGTCAACGGCCAGATCCGCCAATCGGTGCGTGACCGCT	3364
Db	3422	TGCGCAACGTGTGTCAACGTCAACGGCCAGATCCGCCGATCAAGTGGCGGACCGTT	3481
Qy	3365	CGTTGAGTCTGTCAACCGCAACCGCAGAAAGGAGATTCTCCAAACAAGGGCCATGTTG	3424
Db	3482	CGTGGAGTCCGTCAACCGCCACCGCGGAGAGGCGGATGCCACCAACAAGGGCCATGTCG	3541
Qy	3425	CTGACCAATTCGCTGGTGTGTCAACCGTACTGTTGCTGAAGGTGATGAGTCAAGCTG	3484
Db	3542	CGCACCAATTCGCGGTGTGTCAACCGTACTGTTGCTGAAGGTGATGAGTCAAGCTG	3601
Qy	3485	GAGATGCAGTGGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTG	3544
Db	3602	GGACGCGGTGGCCATCATTGAAGGCAATGAAGTGGAGGCCACCATCACCGGCTGTGCG	3661
Qy	3545	ACGCAAAATTCGATCGCTGTGTGTTCTGCTGCAACGAAGGTGGAAGGTGGGACTTGA	3604
Db	3662	ACGGTGTCAATCGACCGCGTTCGTGTGTCGCGCGCCCAAGGTGAGGGCGGCGACCTCA	3721
Qy	3605	TCGTGCTGTTTCCTA	3620
Db	3722	TCGTGCTGTTTCCTA	3737

Search completed: March 24, 2004, 00:13:39  
Job time : 13746 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 17:20:33 ; Search time 1295 Seconds  
(without alignments)  
11878.552 Million cell updates/sec

Title: US-10-045-072-1  
Perfect score: 3621  
Sequence: 1 tggggcgggtagatctg.....tgatcgctgtcttcttaa 3621

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Jan04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002s.\*  
7: Geneseqn2003as.\*  
8: Geneseqn2003bs.\*  
9: Geneseqn2003cs.\*  
10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3621	100.0	3621	3 AAA47533	Sequence
2	3621	100.0	3621	5 AAF32165	AAf32165 Coryneb
3	3621	100.0	3621	6 AAD42046	AAf42046 Coryneb
4	3621	100.0	3621	8 ACA62133	ACA62133 DNA encod
5	3621	100.0	349980	5 AAH68526	AAH68526 C glutami
6	3576.4	98.8	3728	2 AAX24102	Aax24102 C. glutami
7	3420	94.4	3420	5 AAH65730	Aah65730 C glutami
8	3398.8	93.9	3474	6 ABK52832	ABK52832 Coryneb
9	2279.6	63.0	4013	5 AAF87437	AAf87437 Coryneb
10	1927.4	53.2	3423	7 ACA29831	ACA29831 Prokaryot
11	1708	47.2	1719	4 AAF71419	AAf71419 Coryneb
12	1395	38.5	1406	4 AAF71420	AAf71420 Coryneb
13	1361.6	37.6	3381	7 ACA37896	ACA37896 Prokaryot
14	1294.4	35.7	3384	7 ACA40756	ACA40756 Prokaryot
15	1294.4	35.7	110000	4 AAI199682_33	Continuation (34 o
16	1294.4	35.7	110000	4 AAI199683_33	Continuation (34 o
17	1292.2	35.7	3381	7 AAI199683_33	Continuation (34 o
18	1245.6	34.5	32155	9 ADE74252	ADB74252 Mycobacte
19	1033	28.5	1083	4 AAF71418	AAf71418 Coryneb
20	1033	28.5	1083	4 AAF71417	AAf71417 Coryneb
21	907	25.0	939	4 AAF71415	AAf71415 Coryneb
22	907	25.0	939	4 AAF71416	AAf71416 Coryneb
23	681	18.8	3945	6 ABK63659	Rat sequ

24	681	18.8	3945	7 ABT41829	ABt41829 Toxicity
25	681	18.8	3945	9 ADB59092	ADb59092 Toxicity
26	681	18.8	3945	9 ADB52590	ADb52590 Primary r
27	677.4	18.7	4017	6 ABN96532	ABn96532 Gene #303
28	668.4	18.5	4138	4 ABU19577	ABu19577 Drosophil
29	668.4	18.5	4308	4 ABU19575	ABu19575 Drosophil
30	668.4	18.5	4369	4 ABU02789	ABu02789 Drosophil
31	623.4	17.2	4034	4 ABU20985	ABu20985 Drosophil
32	598	16.5	4152	6 ABQ70879	ABq70879 Listeria
33	586	16.2	3468	8 ADB08195	ADb08195 Alloiococ
34	582.4	16.1	110000	6 ADB12064_05	Continuation (6 of
35	582.4	16.1	110000	6 ADB12064_06	Continuation (7 of
36	577.2	15.9	3441	7 ACA36434	ACA36434 Prokaryot
37	577.2	15.9	110000	6 ABA03041_10	Continuation (11 o
38	573.6	15.8	3462	6 AAS01509	AAso1509 Bacillus
39	565	15.6	110000	6 AABQ69245_10	Continuation (11 o
40	565	15.6	110000	6 AABQ67197_09	Continuation (10 o
41	522	14.4	3444	7 ACA22812	ACA22812 Prokaryot
42	514	14.2	3429	7 AAS53072	AAs53072 Enterococ
43	509.2	14.1	3426	7 ACA18474	ACA18474 Prokaryot
44	500.4	13.8	20072	2 AAX13026	AAx13026 Enterococ
45	500.4	13.8	20072	6 ABS98821	ABs98821 Enterococ

## ALIGNMENTS

RESULT 1  
AAA47533  
ID AAA47533 standard; DNA; 3621 BP.  
XX  
AC AAA47533;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
DE Sequence encoding pyruvate carboxylase of C. glutamicum.  
XX  
KW Pyruvate carboxylase; expression; amino acid biosynthesis; lysine;  
XX glutamic acid; oxaloacetate; fermentation; biosynthesis; ds.  
XX Corynebacterium glutamicum.  
XX  
FH Key Location/Qualifiers  
FT CDS  
FT /tag= a  
FT /product= "Pyruvate carboxylase"  
FT /transl\_except= (pos:199..202, aa:Met)

WC2000039305-A1.

XX 06-JUL-2000.

XX 23-DEC-1998; 98WO-US027301.

XX 23-DEC-1998; 98WO-US027301.

XX (SINS/) SINSKEY A. J.

XX (LESS/) LESSARD P. A.

XX (WILL/) WILLIS L. B.

XX Sinskey AJ, Lessard PA, Willis LB;

XX WPI; 2000-465746/40.

XX P-PSDB; AAB01436.

XX Novel polynucleotides encoding Corynebacterium glutamicum pyruvate

XX carboxylase useful for industrial fermentation processes comprises a

XX specific nucleotide sequence.

XX Claim 3; Fig 1; Sipp; English.

XX The pyruvate carboxylase of Corynebacterium glutamicum can be used for

XX producing amino acids, preferably lysine and glutamic acid in industrial





CC The present invention provides the protein and coding sequences of the  
CC Corynebacterium glutamicum pyruvate carboxylase protein. This is an  
CC enzyme in the anaplerotic pathway. It can be used in the replenishment of  
CC oxaloacetate consumed during lysine and glutamic acid production in  
CC industrial fermentation

XX Sequence 3621 BP; 782 A; 1037 C; 1002 G; 800 T; 0 U; 0 Other;

Query Match 100.0%; Score 3621; DB 5; Length 3621;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TGGGGCGGGTGTAGATCCTGGGGGGTTATTTCATTACCTTTGGCTTGAAGTCGTGCAGG	60
Db	1	TGGGGCGGGTGTAGATCCTGGGGGGTTATTTCATTACCTTTGGCTTGAAGTCGTGCAGG	60
Qy	61	TCAGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAACCCGATGTTGATTTGGGGG	120
Db	61	TCAGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAACCCGATGTTGATTTGGGGG	120
Qy	121	AATCGGGGGTTACGATACCTAGGACGAGTCACTGCTATCACCCCTTGGCGGTCTCTTTG	180
Db	121	AATCGGGGGTTACGATACCTAGGACGAGTCACTGCTATCACCCCTTGGCGGTCTCTTTG	180
Qy	181	AAAGGAATAATTAATCTAGTGTGACATCAACATCTTCAACGGTTCACAGCAATTCAAAAG	240
Db	181	AAAGGAATAATTAATCTAGTGTGACATCAACATCTTCAACGGTTCACAGCAATTCAAAAG	240
Qy	241	ATCTTGTGTAGCAACCGCGGAAATCGCGTCCGCTTTCCTGTCAGCACTCGAAACC	300
Db	241	ATCTTGTGTAGCAACCGCGGAAATCGCGTCCGCTTTCCTGTCAGCACTCGAAACC	300
Qy	301	GGTGCAGCAGGTAGTATTTACCCCGGTGAAGATCGGGGATCATTCACCGCTCTTTT	360
Db	301	GGTGCAGCAGGTAGTATTTACCCCGGTGAAGATCGGGGATCATTCACCGCTCTTTT	360
Qy	361	GCTTCTGAAGCTGTCCGATTTGTACCGAAGGCTCACCGTCAAGCGTACCTGGACATC	420
Db	361	GCTTCTGAAGCTGTCCGATTTGTACCGAAGGCTCACCGTCAAGCGTACCTGGACATC	420
Qy	421	GATGAATAATTCGTGTGAGCTTAAAGATGCCATTTACCGGGATACGGC	480
Db	421	GATGAATAATTCGTGTGAGCTTAAAGATGCCATTTACCGGGATACGGC	480
Qy	481	TTCTGTCTGAATAATGCCAGTGTCCCGGAGTGTGCGAATACCGCATTCATTTTAT	540
Db	481	TTCTGTCTGAATAATGCCAGTGTCCCGGAGTGTGCGAATACCGCATTCATTTTAT	540
Qy	541	GGCCCAACCCGAGAGTTCCTGATCTCACCGGTGATAAGTCTCGCGGTAAACCGCGG	600
Db	541	GGCCCAACCCGAGAGTTCCTGATCTCACCGGTGATAAGTCTCGCGGTAAACCGCGG	600
Qy	601	AAGAGGCTGTCTGCCAGTTTGGCGGAATCCACCCGAGCAAAACATCGATGATC	660
Db	601	AAGAGGCTGTCTGCCAGTTTGGCGGAATCCACCCGAGCAAAACATCGATGATC	660
Qy	661	GTTAAAGCGCTGAAGCCAGACTTACCCATCTTTGTGAAGCAGTTGCGGTGTGGC	720
Db	661	GTTAAAGCGCTGAAGCCAGACTTACCCATCTTTGTGAAGCAGTTGCGGTGTGGC	720
Qy	721	GGACGCGGTATGCGTTTGTGCTTACCTGATGAGCTTCGCAATTTAGCAACAGCA	780
Db	721	GGACGCGGTATGCGTTTGTGCTTACCTGATGAGCTTCGCAATTTAGCAACAGCA	780
Qy	781	TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGGCTATATGTCGAACGTGTGAT	840
Db	781	TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGGCTATATGTCGAACGTGTGAT	840
Qy	841	AACCTCAGCATATTAAGTGACATCTTGGCGATCACACTGAGAAAGTTGACCTT	900
Db	841	AACCTCAGCATATTAAGTGACATCTTGGCGATCACACTGAGAAAGTTGACCTT	900
Qy	901	TATGAAGTGAATCTGCTCACTGACGCTCGTCAACCAAAAGTTGTGAAATTTGGCCAGCA	960

Db	901	TATGAAGTGAATCTGCTCACTGACGCTGCTACCAAAAAGTTGTGAAATTTGCCAGCA	960
Qy	961	CAGCATTTGGATCCAGAACTGCGTGTATCGATTGTGCGGATGAGTAAAGTTCTGCCG	1020
Db	961	CAGCATTTGGATCCAGAACTGCGTGTATCGATTGTGCGGATGAGTAAAGTTCTGCCG	1020
Qy	1021	TCCATTGGTTACCAAGCGCGGNAACCGTGAATCTTGGTTCGATGAAGAGCCACAC	1080
Db	1021	TCCATTGGTTACCAAGCGCGGNAACCGTGAATCTTGGTTCGATGAAGAGCCACAC	1080
Qy	1081	GTCTTTCATCGAAATGAACCCAGTATCCAGTGTGAGCACACCGTGTGAGGAAGTCAC	1140
Db	1081	GTCTTTCATCGAAATGAACCCAGTATCCAGTGTGAGCACACCGTGTGAGGAAGTCAC	1140
Qy	1141	GAGTGTGACCTGTGAGGCGGAGTGGCTTGGCTGCTGCTGCTGCACTTGAAGGAATTG	1200
Db	1141	GAGTGTGACCTGTGAGGCGGAGTGGCTTGGCTGCTGCTGCTGCACTTGAAGGAATTG	1200
Qy	1201	CGTCTGACCCCAAGATAAGATCAAGACCCAGCGTGCAGCTGCGAGTCCCGCATCACAC	1260
Db	1201	CGTCTGACCCCAAGATAAGATCAAGACCCAGCGTGCAGCTGCGAGTCCCGCATCACAC	1260
Qy	1261	GAGATCCAAACAAACCGGCTTCCGCCAGATACCGGAACCTATACCGGCTACCGCTACCA	1320
Db	1261	GAGATCCAAACAAACCGGCTTCCGCCAGATACCGGAACCTATACCGGCTACCGCTACCA	1320
Qy	1321	GGCGGAGCTGGCGTTCGCTTACCGTGCAGCTCAGCTCGGTGCGGAAATCACCGCACAC	1380
Db	1321	GGCGGAGCTGGCGTTCGCTTACCGTGCAGCTCAGCTCGGTGCGGAAATCACCGCACAC	1380
Qy	1381	TTTGATCTCAATGCTGTGAAATGACCTGCGTGGTTCGCACTTTGAACTGCTGTGCT	1440
Db	1381	TTTGATCTCAATGCTGTGAAATGACCTGCGTGGTTCGCACTTTGAACTGCTGTGCT	1440
Qy	1441	CGTGACAGCGCGTTCGCTGAGTTCACCGTGTCTGCTGTGCAACCAACATTTGTTTC	1500
Db	1441	CGTGACAGCGCGTTCGCTGAGTTCACCGTGTCTGCTGTGCAACCAACATTTGTTTC	1500
Qy	1501	TTGCGTGGTTCGTGCGGAAAGAGACTTCACTTCCAAAGCGCATCGCCACCGGATTCAT	1560
Db	1501	TTGCGTGGTTCGTGCGGAAAGAGACTTCACTTCCAAAGCGCATCGCCACCGGATTCAT	1560
Qy	1561	GCGGATACCCCGCACTTCTTACCGTCCACTGCTGATGATGAGGAGGACGATCCTG	1620
Db	1561	GCGGATACCCCGCACTTCTTACCGTCCACTGCTGATGATGAGGAGGACGATCCTG	1620
Qy	1621	GATTACTTTGGCAGATGTACCGTGAACAGCCTCATGCTGCTGCTGCTGCAAGGATGTTGCA	1680
Db	1621	GATTACTTTGGCAGATGTACCGTGAACAGCCTCATGCTGCTGCTGCTGCAAGGATGTTGCA	1680
Qy	1681	GCTCCTATCGATAAGCTGCTTAAACATCAAGGATCTGCCACTGCCACCGGTTCCCGTGAC	1740
Db	1681	GCTCCTATCGATAAGCTGCTTAAACATCAAGGATCTGCCACTGCCACCGGTTCCCGTGAC	1740
Qy	1741	GCGCTGAAGCAGCTTGGCCAGCGGTTGCTCGTATCTCCGTGAGCAGGACGCACTG	1800
Db	1741	GCGCTGAAGCAGCTTGGCCAGCGGTTGCTCGTATCTCCGTGAGCAGGACGCACTG	1800
Qy	1801	GCAGTACTGATACCACTTCCCGGATGCAACAGTCTTCTTGTGCGACCCGAGTCCGC	1860
Db	1801	GCAGTACTGATACCACTTCCCGGATGCAACAGTCTTCTTGTGCGACCCGAGTCCGC	1860
Qy	1861	TCATTTCGCACTGAAGCCTGCGGAGGCGCTCGCAAGCTGACTCTCTGAGCTTTTGTTC	1920
Db	1861	TCATTTCGCACTGAAGCCTGCGGAGGCGCTCGCAAGCTGACTCTCTGAGCTTTTGTTC	1920
Qy	1921	GTGAGGCGCTGGGCGGCGACCTACGATGCGGATGCGTTTCTTCTTGTGAGGATCCG	1980
Db	1921	GTGAGGCGCTGGGCGGCGGACCTACGATGCGGATGCGTTTCTTCTTGTGAGGATCCG	1980
Qy	1981	TGGGACGAGCTCGAGAGCTGCGGAGGCGGATGCGGAATGTAACATTCAGATGCTGCTT	2040

Db 1981 TGGACAGGCTCGACGAGCTCGCGAGCGGATGCGGAATGAAACATTCAGATGCTGCTT 2040  
QY 2041 CGCGCGCGACACCGTGGGATACACCCGTAACCGACTCCGTCGCGCGGTTGTT 2100  
Db 2041 CGCGCGCGACACCGTGGGATACACCCGTAACCGACTCCGTCGCGCGGTTGTT 2100  
QY 2101 AAGAAAGCTGACAGCTCGCGGTTGGAATCTTCGCAATCTTCGACGCGCTTAAAGACGCT 2160  
Db 2101 AAGAAAGCTGACAGCTCGCGGTTGGAATCTTCGCAATCTTCGACGCGCTTAAAGACGCT 2160  
QY 2161 TCCAGATGCGTCCAGCAATCGACGAGTCTCGAGACCAACCGCGGTAGCGAGGTG 2220  
Db 2161 TCCAGATGCGTCCAGCAATCGACGAGTCTCGAGACCAACCGCGGTAGCGAGGTG 2220  
QY 2221 GCTATGCTTATTCGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCTGGATTAC 2280  
Db 2221 GCTATGCTTATTCGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCTGGATTAC 2280  
QY 2281 TACCTAAGATGCGAGGAGATCGTCAAGTCTGCGGCTCACATCTTGGCCATTAAAGAT 2340  
Db 2281 TACCTAAGATGCGAGGAGATCGTCAAGTCTGCGGCTCACATCTTGGCCATTAAAGAT 2340  
QY 2341 ATGCTGCTCTGCTTCGCGCCAGCTGCGGTAAACCAAGCTGCTCAACGCACTGCGCGGTAA 2400  
Db 2341 ATGCTGCTCTGCTTCGCGCCAGCTGCGGTAAACCAAGCTGCTCAACGCACTGCGCGGTAA 2400  
QY 2401 TTGCATCTGCCAGTGCAGTGCAACCCACGACACTGCGGGTGGCCAGCTGGCACTTAC 2460  
Db 2401 TTGCATCTGCCAGTGCAGTGCAACCCACGACACTGCGGGTGGCCAGCTGGCACTTAC 2460  
QY 2461 TTTGCTGAGCTCAAGCTGCTGAGATGCTGTTGACGCTGCTTCGCGCACCACTGCTGCG 2520  
Db 2461 TTTGCTGAGCTCAAGCTGCTGAGATGCTGTTGACGCTGCTTCGCGCACCACTGCTGCG 2520  
QY 2521 ACCAGCTCCAGCAATCCCTGCTGCGCAATTTGCTGCAATTCGCGCAACCGCTGCGGAT 2580  
Db 2521 ACCAGCTCCAGCAATCCCTGCTGCGCAATTTGCTGCAATTCGCGCAACCGCTGCGGAT 2580  
QY 2581 ACCGGTTTGAGCTCGAGCTGTTTCTGACCTCGAGCGGTACTGGGAAGCAGTGGCGGA 2640  
Db 2581 ACCGGTTTGAGCTCGAGCTGTTTCTGACCTCGAGCGGTACTGGGAAGCAGTGGCGGA 2640  
QY 2641 CTGTACCTGCCATTTGAGCTGGAACCCAGCCCAACCGGTGCGCTTACCGCCAGAA 2700  
Db 2641 CTGTACCTGCCATTTGAGCTGGAACCCAGCCCAACCGGTGCGCTTACCGCCAGAA 2700  
QY 2701 ATCCAGCGGACAGCTGCTCAACCTGCTGCAAGGCCACCGCACTGGGCTTGGCGAT 2760  
Db 2701 ATCCAGCGGACAGCTGCTCAACCTGCTGCAAGGCCACCGCACTGGGCTTGGCGAT 2760  
QY 2761 CGTTTCGAACCTCATCGAAGCAACTACCGACCGGTTAATGAGATGCTGGGACGCCAAC 2820  
Db 2761 CGTTTCGAACCTCATCGAAGCAACTACCGACCGGTTAATGAGATGCTGGGACGCCAAC 2820  
QY 2821 AAGGTCACCCATCTCCAGGTTTGGGACCTCGCACTCGCACTCGGCTTGGGCGGAT 2880  
Db 2821 AAGGTCACCCATCTCCAGGTTTGGGACCTCGCACTCGCACTCGGCTTGGGCGGAT 2880  
QY 2881 GTGGATCCAGCAGACTTTGCTGCGGATCCCAAAAGTACGACATCCCAAGATCTGTATC 2940  
Db 2881 GTGGATCCAGCAGACTTTGCTGCGGATCCCAAAAGTACGACATCCCAAGATCTGTATC 2940  
QY 2941 GGGTTCTGCGCGGAGCTTGGTAACCTTCAGGTGGCTGGCGAGCCACTGGCACC 3000  
Db 2941 GGGTTCTGCGCGGAGCTTGGTAACCTTCAGGTGGCTGGCGAGCCACTGGCACC 3000  
QY 3001 CGCGCACTGGAAGCGGCTCCAGGCAAGGCACTCTCTGAGCGGAGTTCTTGAAGAGAG 3060  
Db 3001 CGCGCACTGGAAGCGGCTCCAGGCAAGGCACTCTCTGAGCGGAGTTCTTGAAGAGAG 3060  
QY 3061 CAGGCGCACTCGAGCTGATGATTCAGAGAACCTGCAATAGCTCAACCGCTGCTG 3120  
Db 3061 CAGGCGCACTCGAGCTGATGATTCAGAGAACCTGCAATAGCTCAACCGCTGCTG 3120

QY 3121 TTCCGAAGCCAAACCGAGAGTTCTCTGAGCAACCGTCCCGCTTCGCAACACCTCTGCG 3180  
Db 3121 TTCCGAAGCCAAACCGAGAGTTCTCTGAGCAACCGTCCCGCTTCGCAACACCTCTGCG 3180  
QY 3181 CTGGATGATCGTGAATTTCTTACGGCTGCTGAAGCGCGGAGACTTTTGTATCGGCTG 3240  
Db 3181 CTGGATGATCGTGAATTTCTTACGGCTGCTGAAGCGCGGAGACTTTTGTATCGGCTG 3240  
QY 3241 CCAGATGCGCCACCCCACTGCTTTGTCGCTGATGCGATCTCTGAGCCAGAGATAAG 3300  
Db 3241 CCAGATGCGCCACCCCACTGCTTTGTCGCTGATGCGATCTCTGAGCCAGAGATAAG 3300  
QY 3301 GGTATGCCAATGTTGGCCAAAGTCAACCGCCAGATCCGCCCAATGCGTGTGCGTAC 3360  
Db 3301 GGTATGCCAATGTTGGCCAAAGTCAACCGCCAGATCCGCCCAATGCGTGTGCGTAC 3360  
QY 3361 CGCTCCGTTGAGTCTGTCAACCGCAACCGCAGAAAGGAGATTTCTCCAAACAGGGCCAT 3420  
Db 3361 CGCTCCGTTGAGTCTGTCAACCGCAACCGCAGAAAGGAGATTTCTCCAAACAGGGCCAT 3420  
QY 3421 GTTGTGCAACATTCGCTGGTGTGTACCGTGTGTTGCTGAGTGTGAGGTGAGTCAAG 3480  
Db 3421 GTTGTGCAACATTCGCTGGTGTGTACCGTGTGTTGCTGAGTGTGAGGTGAGTCAAG 3480  
QY 3481 GCTGAGATGCGATGCGCAATCATCGAGCTATGAGATGGAAGCAACAAATCACTGCTTCT 3540  
Db 3481 GCTGAGATGCGATGCGCAATCATCGAGCTATGAGATGGAAGCAACAAATCACTGCTTCT 3540  
QY 3541 GTTGCAGCAAAATCGATCGGTTGTTGCTGCTGCAACGAGGTGGAAGTGGCGAC 3600  
Db 3541 GTTGCAGCAAAATCGATCGGTTGTTGCTGCTGCAACGAGGTGGAAGTGGCGAC 3600  
QY 3601 TTGATCGCTGCTGTTTCTTAA 3621  
Db 3601 TTGATCGCTGCTGTTTCTTAA 3621

## RESULT 3

AAD42046  
ID AAD42046 standard; DNA; 3621 BP.  
XX  
AC AAD42046;  
XX  
DT 04-NOV-2002 (first entry)  
XX  
DE Corynebacterium glutamicum pyruvate carboxylase DNA.  
XX  
KW Pyruvate carboxylase; anaplerotic enzyme; industrial fermentation;  
XX oxaloacetate; growth; enzyme; gene; ds.  
XX  
OS Corynebacterium glutamicum.

## Key Location/Qualifiers

CDS  
199..3621  
/\*tag= a  
/product= "Pyruvate carboxylase protein"  
/transl\_except= (pos:199..201, aa:Met)

US6403351-B1.

11-JUN-2002.

03-OCT-2000; 2000US-00677575.

23-DEC-1998; 98US-00220081.

(ARCH ) ARCHER-DANIELS MIDLAND CO.

Sinskey AJ, Lessard PA, Willis LB;

WPI; 2002-536037/57.

P-PSDB; AAE25601.





1861 TCATTGCACTGAAGCCTGCGGCAGAGCCGCTCGCAAGAGTGACTCTGAGCTTTTGTCC 1920  
1921 GTGAGGCTTGGGCGGCGGAGCCTACGATGTGCGATGGTTTCCCTTTGAGGATCCG 1980  
1921 GTGAGGCTTGGGCGGCGGAGCCTACGATGTGCGATGGTTTCCCTTTGAGGATCCG 1980  
1981 TGGGACAGGCTCGACGAGCTCGCGAGCGGAGTGCAGATGTAAACATTCAGATGTGCTT 2040  
1981 TGGGACAGGCTCGACGAGCTCGCGAGCGGAGTGCAGATGTAAACATTCAGATGTGCTT 2040  
2041 CGCGCCGACACCGTGGGATACACCCGACCGTACCGACTCCGCTCCCGCGGTTGTT 2100  
2041 CGCGCCGACACCGTGGGATACACCCGACCGTACCGACTCCGCTCCCGCGGTTGTT 2100  
2101 AAGGAAGCTGCGAGCTCGCGGTGAGCATCTTCGCGCATCTTCGAGCGGCTTAACGAGTC 2160  
2101 AAGGAAGCTGCGAGCTCGCGGTGAGCATCTTCGCGCATCTTCGAGCGGCTTAACGAGTC 2160  
2161 TCCGAGATGCTCCAGCATCGAGCAGCTCTGAGAGCCACACCGGCTAGCCGAGGTG 2220  
2161 TCCGAGATGCTCCAGCATCGAGCAGCTCTGAGAGCCACACCGGCTAGCCGAGGTG 2220  
2221 GCTATGGCTTAATCTGCTGATCTCTGATCCAAATGAAAGCTCTACACCTCGATTAC 2280  
2221 GCTATGGCTTAATCTGCTGATCTCTGATCCAAATGAAAGCTCTACACCTCGATTAC 2280  
2281 TACCTAAAGATGCGAGAGAGATCTCAAGTCTGCGCTCACATCTTGCGCATTAAGGAT 2340  
2281 TACCTAAAGATGCGAGAGAGATCTCAAGTCTGCGCTCACATCTTGCGCATTAAGGAT 2340  
2341 ATGGCTGCTCTGCTTCGCGCAGCTCGGTAAACCAAGCTGGTCAACGCTGCGCGTGA 2400  
2341 ATGGCTGCTCTGCTTCGCGCAGCTCGGTAAACCAAGCTGGTCAACGCTGCGCGTGA 2400  
2401 TTGCAATCCAGTGCAGTGCACACCGACGACATCTGCGGTGCGCAGCTGCGACCTAC 2460  
2401 TTGCAATCCAGTGCAGTGCACACCGACGACATCTGCGGTGCGCAGCTGCGACCTAC 2460  
2461 TTGCTGCAAGCTCAAGCTGGTGCAGATGCTGTGACGGTGTCTCGCGACCACTGTCTGC 2520  
2461 TTGCTGCAAGCTCAAGCTGGTGCAGATGCTGTGACGGTGTCTCGCGACCACTGTCTGC 2520  
2521 ACCACTCCAGCATCCCTGTCTGCAATGTTGTGCTGATTCGCGCACACCGCTGCGCAT 2580  
2521 ACCACTCCAGCATCCCTGTCTGCAATGTTGTGCTGATTCGCGCACACCGCTGCGCAT 2580  
2581 ACCGGTTTGCAGCTCGAGGCTGTTCTGACCTCGAGCGCTACTGCGAAGCAGTGGCGGA 2640  
2581 ACCGGTTTGCAGCTCGAGGCTGTTCTGACCTCGAGCGCTACTGCGAAGCAGTGGCGGA 2640  
2641 CTGTACTGCTCAATTTGAGTCTGGAACCCCGAGCCCAACCGGTGCGGTCTACCGCCAGAA 2700  
2641 CTGTACTGCTCAATTTGAGTCTGGAACCCCGAGCCCAACCGGTGCGGTCTACCGCCAGAA 2700  
2701 ATCCGAGCGGACAGTGTCCACCTGCTGCAAGCCGACAGCCGACCTGGGCTTGGCAT 2760  
2701 ATCCGAGCGGACAGTGTCCACCTGCTGCAAGCCGACAGCCGACCTGGGCTTGGCAT 2760  
2761 CGTTTCGAACTCATCGAAGCACTACGAGCGGTTAATGAGATGCTGGAGCGCCCAAC 2820  
2761 CGTTTCGAACTCATCGAAGCACTACGAGCGGTTAATGAGATGCTGGAGCGCCCAAC 2820  
2821 AAGGTACCCCATCTCCAGGTTGTTGGAGCTCGCATCTCACCTCGTGTGGTGGGT 2880  
2821 AAGGTACCCCATCTCCAGGTTGTTGGAGCTCGCATCTCACCTCGTGTGGTGGGT 2880  
2881 GTGATCCAGCAGACTTTGCTGCGGATCCAAAGTACGACATCCAGACTCTGTCTATC 2940  
2881 GTGATCCAGCAGACTTTGCTGCGGATCCAAAGTACGACATCCAGACTCTGTCTATC 2940  
2941 GCGTTCCTGCGGCGAGCTTGGTAAACCTTCAGGTTGGTGGCGAGACCTGCGCAC 3000  
2941 GCGTTCCTGCGGCGAGCTTGGTAAACCTTCAGGTTGGTGGCGAGACCTGCGCAC 3000

3001 CCGCACTGGAAGCGCGCTCCGAAGCGAGGCACTCTGACGGAAGTTCTTGGAGAGAG 3060  
3001 CCGCACTGGAAGCGCGCTCCGAAGCGAGGCACTCTGACGGAAGTTCTTGGAGAGAG 3060  
3061 CAGGCGCACTCGACGCTGATGATTCCAAAGAAAGTTCGAATAGCTCAACCGCTGCTG 3120  
3061 CAGGCGCACTCGACGCTGATGATTCCAAAGAAAGTTCGAATAGCTCAACCGCTGCTG 3120  
3121 TTCCGAAGCAACCGAAGAGTTCCTGAGCAACCGTTCGCCGCTTCGGAACACCTCTCG 3180  
3121 TTCCGAAGCAACCGAAGAGTTCCTGAGCAACCGTTCGCCGCTTCGGAACACCTCTCG 3180  
3181 CTGGATGATCGGAATCTCTTACCGCTGCTGGAAGCGCGGAGCTTTGATCCGCTG 3240  
3181 CTGGATGATCGGAATCTCTTACCGCTGCTGGAAGCGCGGAGCTTTGATCCGCTG 3240  
3241 CCAGATGTGCGCAACCGCACTGCTTGTTCGCTGGAATGCGATCTCTGAGCCAGACGATAG 3300  
3241 CCAGATGTGCGCAACCGCACTGCTTGTTCGCTGGAATGCGATCTCTGAGCCAGACGATAG 3300  
3301 GGTATGCGCAATGTTGTGCGCAACCGTCAACGCGCAGATCCGCCCAATGCGTGTGCGTGA 3360  
3301 GGTATGCGCAATGTTGTGCGCAACCGTCAACGCGCAGATCCGCCCAATGCGTGTGCGTGA 3360  
3361 CGTCCGTTGAGTCTGTCAACCGCAACCGCAGAAAGGCGAGATTCCTCAACGAGGCGCAT 3420  
3361 CGTCCGTTGAGTCTGTCAACCGCAACCGCAGAAAGGCGAGATTCCTCAACGAGGCGCAT 3420  
3421 GTTGTGCAACCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480  
3421 GTTGTGCAACCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480  
3481 GCTGAGATGCACTGCAATCATCGAGGCTATGAAGATGGAAGCAACATCATCTGCTTCT 3540  
3481 GCTGAGATGCACTGCAATCATCGAGGCTATGAAGATGGAAGCAACATCATCTGCTTCT 3540  
3541 GTTGAAGCAAAATCGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600  
3541 GTTGAAGCAAAATCGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600  
3601 TTGATGCTGCTGCTTTCTTAA 3621  
3601 TTGATGCTGCTGCTTTCTTAA 3621

## RESULT 4

ACA62133

ID ACA62133 standard; DNA; 3621 BP.

XX ACA62133;

XX 25-AUG-2003 (first entry)

XX DNA encoding Corynebacterium glutamicum pyruvate carboxylase.

XX Pyruvate carboxylase; gene; anaplerotic enzyme; oxaloacetate;  
XX biosynthesis; growth; lysine production; glutamic acid production;  
XX industrial fermentation; gene; ds.

XX Corynebacterium glutamicum.

XX Key Location/Qualifiers  
XX CDS 199..3621  
XX /\*tag= a  
XX /product= "Pyruvate carboxylase"

XX

XX US2003027305-A1.

XX 06-FEB-2003.

XX 15-JAN-2002; 2002US-00045072.

XX

PR 23-DEC-1998; 98US-00220081.  
 PR 03-OCT-2000; 2000US-00677575.  
 PA (ARCH) ARCHER-DANIELS MIDLAND CO.  
 XX Sinekey AJ, Lessard PA, Willis LB;  
 XX P-PSDB; ABU10425.  
 XX WPI; 2003-479542/58.  
 DR P-PSDB; ABU10425.  
 XX  
 PT New pyruvate carboxylase from *Corynebacterium glutamicum*, useful as an  
 PT anaplerotic enzyme replenishing oxaloacetate consumed for biosynthesis  
 PT during growth, or for lysine or glutamic acid production in industrial  
 PT fermentations.  
 XX  
 PS Claim 2; Page 12-16; 29pp; English.  
 CC The invention describes a new isolated pyruvate carboxylase polypeptide  
 CC having an amino acid sequence at least 95% identical to a sequence  
 CC comprising 1140 amino acids from *Corynebacterium glutamicum*, or the  
 CC complete amino acid sequence encoded by the cosmid clone deposited with  
 CC the American Type Culture Collection. The polypeptide is useful as an  
 CC anaplerotic enzyme replenishing oxaloacetate consumed for biosynthesis  
 CC during growth. The polypeptide is also useful for lysine or glutamic acid  
 CC production in industrial fermentations. This sequence encodes  
 CC *Corynebacterium glutamicum* pyruvate carboxylase  
 XX  
 SQ Sequence 3621 BP; 782 A; 1037 C; 1002 G; 800 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 3621; DB 8; Length 3621;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGGGGGGGTTAGATCTGGGGGGTTATTTCAATTCATCTTGGTGAAGTCGTCAGG 60  
 Db 1 TGGGGGGGGTTAGATCTGGGGGGTTATTTCAATTCATCTTGGTGAAGTCGTCAGG 60

Qy 61 TCAGGGAGGTGTTCCCGGAAACATTGAGAGGAAACAAACACCGATGTTGATTCGGGG 120  
 Db 61 TCAGGGAGGTGTTCCCGGAAACATTGAGAGGAAACAAACACCGATGTTGATTCGGGG 120

Qy 121 AATCGGGGTTACGATAGGACGAGTGAATTCACCTTCCGCGTCTCTTTGTTG 180  
 Db 121 AATCGGGGTTACGATAGGACGAGTGAATTCACCTTCCGCGTCTCTTTGTTG 180

Qy 181 AAGGAATTAATCTCTAGTGTGACTCACATCTTCCACGCTTCCAGCATTCAAAAG 240  
 Db 181 AAGGAATTAATCTCTAGTGTGACTCACATCTTCCACGCTTCCAGCATTCAAAAG 240

Qy 241 ATCTTGTAGCAACACCGCGGAAATCCGCGTCCGTTTCCGTCAGCACTTCGAAACC 300  
 Db 241 ATCTTGTAGCAACACCGCGGAAATCCGCGTCCGTTTCCGTCAGCACTTCGAAACC 300

Qy 301 GGTCCAGCCAGTGTATTTACCCCGTGAGATCGGGGATTCATCCACGCTCTTTT 360  
 Db 301 GGTCCAGCCAGTGTATTTACCCCGTGAGATCGGGGATTCATCCACGCTCTTTT 360

Qy 361 GCTTCTGAAGTGTCCGATTTGGTACCGAAGGCTCACAGTCAAGCGTACCTGGACATC 420  
 Db 361 GCTTCTGAAGTGTCCGATTTGGTACCGAAGGCTCACAGTCAAGCGTACCTGGACATC 420

Qy 421 GATCAATTAATCGTGCAGTGAATAAGTTAAAGCAGATGCCATTTACCCGGATACGGC 480  
 Db 421 GATCAATTAATCGTGCAGTGAATAAGTTAAAGCAGATGCCATTTACCCGGATACGGC 480

Qy 481 TTCTGTCTGAAATATGCCAGTCTTCCCGGAGTGTGCGGAAACCGCATTTACTTTTAT 540  
 Db 481 TTCTGTCTGAAATATGCCAGTCTTCCCGGAGTGTGCGGAAACCGCATTTACTTTTAT 540

Qy 541 GGCCCAACCCAGAGTCTTGTATCTCACCGGTGATAGTCTCGCGGTTAACCGCGCG 600  
 Db 541 GGCCCAACCCAGAGTCTTGTATCTCACCGGTGATAGTCTCGCGGTTAACCGCGCG 600

Qy 601 AAGAGAGCTGTGCTGCCAGTTTGGCGGAATCCACCCGAGCAAAACATCGATGATC 660  
 Db 601 AAGAGAGCTGTGCTGCCAGTTTGGCGGAATCCACCCGAGCAAAACATCGATGATC 660

Qy 661 GTTAAAGAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAGCGAGTTGCCGTTGGC 720  
 Db 661 GTTAAAGAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAGCGAGTTGCCGTTGGC 720

Qy 721 GGAGCGGTATGCGTTTGTGCTTACCTGATGAGCTTCGCAAAATAGCAACAGAGCA 780  
 Db 721 GGAGCGGTATGCGTTTGTGCTTACCTGATGAGCTTCGCAAAATAGCAACAGAGCA 780

Qy 781 TCTCGTGAAGCTGAAGCGGCTTTTCGCGCATGGCGCGTATATGTCGAACGTGCTGATT 840  
 Db 781 TCTCGTGAAGCTGAAGCGGCTTTTCGCGCATGGCGCGTATATGTCGAACGTGCTGATT 840

Qy 841 AACCTCAGCATATTGAAGTGCAGATCTTGGCGATCACAATGGAGAGTTGTACACCTT 900  
 Db 841 AACCTCAGCATATTGAAGTGCAGATCTTGGCGATCACAATGGAGAGTTGTACACCTT 900

Qy 901 TATGAACGTGACTGCTCACTGACGCTCGTCACCAAAAAGTTGTGCAAAATTCGCCAGCA 960  
 Db 901 TATGAACGTGACTGCTCACTGACGCTCGTCACCAAAAAGTTGTGCAAAATTCGCCAGCA 960

Qy 961 CAGCATTTGGATCCAGACTGCGTATGCGATTTGTGCGGATCAGTAAGTTCTGCCG 1020  
 Db 961 CAGCATTTGGATCCAGACTGCGTATGCGATTTGTGCGGATCAGTAAGTTCTGCCG 1020

Qy 1021 TCCATTGGTTACCAAGGCGCGGAAACCGTGAATTTCTTGGTTCATGAAAGGCAACAC 1080  
 Db 1021 TCCATTGGTTACCAAGGCGCGGAAACCGTGAATTTCTTGGTTCATGAAAGGCAACAC 1080

Qy 1081 GTCTTCATCGAAATGAACCCAGCTATCCAGTTTGTGAGCACHACCGTGAAGTCAAC 1140  
 Db 1081 GTCTTCATCGAAATGAACCCAGCTATCCAGTTTGTGAGCACHACCGTGAAGTCAAC 1140

Qy 1141 GAGTGGACCTGTGTGAAGGCGCAGATCGCTTGGCTGTGCTGCAACCTTGAAGGAATTG 1200  
 Db 1141 GAGTGGACCTGTGTGAAGGCGCAGATCGCTTGGCTGTGCTGCAACCTTGAAGGAATTG 1200

Qy 1201 GGTTCAGCCCAAGATAGATCAAGACCAAGTGTGAGCACHACCGTGAAGTCAAC 1260  
 Db 1201 GGTTCAGCCCAAGATAGATCAAGACCAAGTGTGAGCACHACCGTGAAGTCAAC 1260

Qy 1261 GAAATCCCAACAAACGGCTTCCGCGCAGATACCGAACTATCACCGGTACCGCTCACCA 1320  
 Db 1261 GAAATCCCAACAAACGGCTTCCGCGCAGATACCGAACTATCACCGGTACCGCTCACCA 1320

Qy 1321 GGGGAGCTGCGTGTGCTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1380  
 Db 1321 GGGGAGCTGCGTGTGCTTTGACGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1380

Qy 1381 TTTGACTCCATGCTGTGAAATGACCTGCGTGGTTCCGACTTTTGAACCTGCTGTTGCT 1440  
 Db 1381 TTTGACTCCATGCTGTGAAATGACCTGCGTGGTTCCGACTTTTGAACCTGCTGTTGCT 1440

Qy 1441 CGTGCACAGCGCGGTTGGCTGAGTTACCGTGTCTGGTGTGCAACCAACATTTGGTTTC 1500  
 Db 1441 CGTGCACAGCGCGGTTGGCTGAGTTACCGTGTCTGGTGTGCAACCAACATTTGGTTTC 1500

Qy 1501 TTGCGTGTGTTGTCGCGGAAGAGGACTTCACTTCCAAAGCGCATCGCACCGGATTCATT 1560  
 Db 1501 TTGCGTGTGTTGTCGCGGAAGAGGACTTCACTTCCAAAGCGCATCGCACCGGATTCATT 1560

Qy 1561 GCCGATACCCCGCACTCTTTCAGGCTTCCACTGCTGATGATGAGGAGGAGCGCATCTG 1620  
 Db 1561 GCCGATACCCCGCACTCTTTCAGGCTTCCACTGCTGATGATGAGGAGGAGCGCATCTG 1620

Qy 1621 GATTACTTGGCAGATGTCAACCGTGAACAGCTCATGCTGTGCTGTCCAAAGGATGTTGCA 1680  
 Db 1621 GATTACTTGGCAGATGTCAACCGTGAACAGCTCATGCTGTGCTGTCCAAAGGATGTTGCA 1680

Qy 1681 GCTCCTATCGATAAGCTGCTTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740

Db 1681 GCTCTTATGATAAGCTGCTTAACATCAAGGATCTGCACTGCGAGGGTTCCGCTGAC 1740  
QY 1741 CGCCTGAAGAGCTTGGCCAGCGCGGTTGCTGCTGATCTCGTGAGAGCAACCTG 1800  
Db 1741 CGCCTGAAGAGCTTGGCCAGCGCGGTTGCTGCTGATCTCGTGAGAGCAACCTG 1800  
QY 1801 GCAGTTACTGATACACCTTCCGCGATGACACCACTGCTTCTGCTGCGACCGAGTCCGC 1860  
Db 1801 GCAGTTACTGATACACCTTCCGCGATGACACCACTGCTTCTGCTGCGACCGAGTCCGC 1860  
QY 1861 TCATTGCGACTGAAGCTTGGGAGAGGCGGTGCGAAAGCTGACTCTGAGCTTTGTC 1920  
Db 1861 TCATTGCGACTGAAGCTTGGGAGAGGCGGTGCGAAAGCTGACTCTGAGCTTTGTC 1920  
QY 1921 GTGGAGGCTTGGGCGGCGGAGCTTACGATGTGGGATGCTTCTCTTTGAGGATCCG 1980  
Db 1921 GTGGAGGCTTGGGCGGCGGAGCTTACGATGTGGGATGCTTCTCTTTGAGGATCCG 1980  
QY 1981 TGGGACAGGCTGACGAGCTGGCGAGGCGATGCCGAATGTAAACATTGAGATGCTGTT 2040  
Db 1981 TGGGACAGGCTGACGAGCTGGCGAGGCGATGCCGAATGTAAACATTGAGATGCTGTT 2040  
QY 2041 CGCGGCGCAACACCGTGGGATACACCGTACCAGACTCCGCTGCGCGGCTTTGTT 2100  
Db 2041 CGCGGCGCAACACCGTGGGATACACCGTACCAGACTCCGCTGCGCGGCTTTGTT 2100  
QY 2101 AAGGAAGCTGCCAGCTCCGCGTGGACATCTTCCGATCTTCGACGCGTTAAACGACGTC 2160  
Db 2101 AAGGAAGCTGCCAGCTCCGCGTGGACATCTTCCGATCTTCGACGCGTTAAACGACGTC 2160  
QY 2161 TCCAGATCGCTCCAGCAATCGACGAGCTCTGGAGACCAACACCGTAGCGAGGTC 2220  
Db 2161 TCCAGATCGCTCCAGCAATCGACGAGCTCTGGAGACCAACACCGTAGCGAGGTC 2220  
QY 2221 GCTATGGCTTATCTGGTATCTCTGATCAAAATGAAAGCTCTACACCTGGAATTAC 2280  
Db 2221 GCTATGGCTTATCTGGTATCTCTGATCAAAATGAAAGCTCTACACCTGGAATTAC 2280  
QY 2281 TACCTAAGATGGCAGAGAGATCTGATGCTGGGCTCACATCTTGGCCATTGAAGAT 2340  
Db 2281 TACCTAAGATGGCAGAGAGATCTGATGCTGGGCTCACATCTTGGCCATTGAAGAT 2340  
QY 2341 ATGGCTGGTCTGCTTCCGCGAGCTGGGTAAACCAAGCTGCTACCGCACTGGCGGTGAA 2400  
Db 2341 ATGGCTGGTCTGCTTCCGCGAGCTGGGTAAACCAAGCTGCTACCGCACTGGCGGTGAA 2400  
QY 2401 TTCGATCTGCGAGTGCAGCTGCACACCCAGACACTGCGGGTGGCCAGTGGCAACCTAC 2460  
Db 2401 TTCGATCTGCGAGTGCAGCTGCACACCCAGACACTGCGGGTGGCCAGTGGCAACCTAC 2460  
QY 2461 TTTGCTGCAAGCTCAAGCTGGTGAGATGCTGTTGAGCGGTTCGCGACCACTGCTGGC 2520  
Db 2461 TTTGCTGCAAGCTCAAGCTGGTGAGATGCTGTTGAGCGGTTCGCGACCACTGCTGGC 2520  
QY 2521 ACCACTCCAGCCATCCCTGCTGCGCATTTGCTGATTCGCGACACCGGTGCGCAT 2580  
Db 2521 ACCACTCCAGCCATCCCTGCTGCGCATTTGCTGATTCGCGACACCGGTGCGCAT 2580  
QY 2581 ACCGGTTTGAGCTCGAGCTGTTCTGACCTCGAGCGCTACTGGGAAGCAGTGGCGGA 2640  
Db 2581 ACCGGTTTGAGCTCGAGCTGTTCTGACCTCGAGCGCTACTGGGAAGCAGTGGCGGA 2640  
QY 2641 CTGTACTGCGCATTTGAGTCTGGAACCCCGAGCCCAACCGGTGCGGTCTACCGCACGAA 2700  
Db 2641 CTGTACTGCGCATTTGAGTCTGGAACCCCGAGCCCAACCGGTGCGGTCTACCGCACGAA 2700  
QY 2701 ATCCAGGCGGACAGTTGCTCAACCTGCTGACAGGCGCCAGCTGGGCTTGGCAT 2760  
Db 2701 ATCCAGGCGGACAGTTGCTCAACCTGCTGACAGGCGCCAGCTGGGCTTGGCAT 2760  
QY 2761 CGTTTCGAATCTCAAGACCAACTACGAGCGGTTAATGAGATGCTGGGACGCGCAAC 2820

Db 2761 CGTTTCGAATCTCAAGACCAACTACGAGCGGTTAATGAGATGCTGGGACGCGCAAC 2820  
QY 2821 AAGTCAACCCATCTCCAAAGTTGTTGGGAGCTCGACTCGACTCCACTCGTTGTTGGGTT 2880  
Db 2821 AAGTCAACCCATCTCCAAAGTTGTTGGGAGCTCGACTCGACTCGTTGTTGGGTT 2880  
QY 2881 GTGATCCAGCAGACTTGTGCTGCGATCCACAAAAGTACGACATCCAGACTCTGTATC 2940  
Db 2881 GTGATCCAGCAGACTTGTGCTGCGATCCACAAAAGTACGACATCCAGACTCTGTATC 2940  
QY 2941 GCGTTCGTCGCGGAGCTTGTAAACCTTCAGTGTGTCGCGCAGACCACTGCGGAC 3000  
Db 2941 GCGTTCGTCGCGGAGCTTGTAAACCTTCAGTGTGTCGCGCAGACCACTGCGGAC 3000  
QY 3001 CGGCACTGAAAGCGCGCTCCGAAGCAAGGCACTCTGACGGAAGTTCCTGAGGAGAG 3060  
Db 3001 CGGCACTGAAAGCGCGCTCCGAAGCAAGGCACTCTGACGGAAGTTCCTGAGGAGAG 3060  
QY 3061 CAGCGCACTCGACGCTGATGATTCGAAGAACTGCGCAATAGCCCTCAACCGCTGCTG 3120  
Db 3061 CAGCGCACTCGACGCTGATGATTCGAAGAACTGCGCAATAGCCCTCAACCGCTGCTG 3120  
QY 3121 TTCCCGAAGCAACCGAAGGTTCTTCGAGCACCGTCGCGCTTCGCGCAACCTCTGG 3180  
Db 3121 TTCCCGAAGCAACCGAAGGTTCTTCGAGCACCGTCGCGCTTCGCGCAACCTCTGG 3180  
QY 3181 CTGATGATCGTGAATTTCTTACGCTTGTTCGAAGCGCGGAGACTTTTGATTCGCTG 3240  
Db 3181 CTGATGATCGTGAATTTCTTACGCTTGTTCGAAGCGCGGAGACTTTTGATTCGCTG 3240  
QY 3241 CCAGATGCGCACCCACCTGCTTGTTCGCTGGATGCGATCTCTGAGCCAGACGATAAG 3300  
Db 3241 CCAGATGCGCACCCACCTGCTTGTTCGCTGGATGCGATCTCTGAGCCAGACGATAAG 3300  
QY 3301 GGTATGCGCAATTTGTGGCCAACTGCAAGCGCCAGATCCGCCCAATGCGTGTGCGTAC 3360  
Db 3301 GGTATGCGCAATTTGTGGCCAACTGCAAGCGCCAGATCCGCCCAATGCGTGTGCGTAC 3360  
QY 3361 CGCTCGCTTGTGCTGTCACCGCAACCGGAGAAAGGAGACTTCCTCAACAGGCGCAT 3420  
Db 3361 CGCTCGCTTGTGCTGTCACCGCAACCGGAGAAAGGAGACTTCCTCAACAGGCGCAT 3420  
QY 3421 GTTGCTGCACTTCTGCTGCTGTTGTTCAGCTGCTGTTGCTGAGGTTGATGAGTCAAG 3480  
Db 3421 GTTGCTGCACTTCTGCTGCTGTTGTTCAGCTGCTGTTGCTGAGGTTGATGAGTCAAG 3480  
QY 3481 GCTGGAGATGCTGCAATCATCGAGGCTATGAAAGTGAAGCAACATCACTGCTTCT 3540  
Db 3481 GCTGGAGATGCTGCAATCATCGAGGCTATGAAAGTGAAGCAACATCACTGCTTCT 3540  
QY 3541 GTTGACGGCAAAATCGATCGCTTGTGTTCTGCTGCAACGAAAGTGAAGGTCGCGAC 3600  
Db 3541 GTTGACGGCAAAATCGATCGCTTGTGTTCTGCTGCAACGAAAGTGAAGGTCGCGAC 3600  
QY 3601 TTGATCGTCTGTTTCTTAA 3621  
Db 3601 TTGATCGTCTGTTTCTTAA 3621

## RESULT 5

AAH68526 standard; DNA; 349980 BP.

XX AAH68526;

XX 26-SEP-2001 (first entry)

XX C glutamicum coding sequence fragment SEQ ID NO: 7061.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis; ds.

XX Corynebacterium glutamicum.







AC AAX24102;  
XX  
DT 30-JUN-1999 (first entry)  
XX  
DE C. glutamicum pyruvate carboxylase genomic DNA.  
XX  
KW Pyruvate carboxylase; amino acid production; lysine production;  
KW threonine production; homoserine production; glutamate production;  
KW arginine production; feed additive; condiment; pharmaceutical;  
KW fine chemical; ss.  
XX  
OS Corynebacterium glutamicum.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 165..3587  
FT /\*tag= a  
FT /product= "pyruvate carboxylase"  
XX  
XX DE19831609-Al.  
XX  
XX 15-APR-1999.  
XX  
XX 14-JUL-1998; 98DB-01031609.  
XX  
XX 04-OCT-1997; 97DB-01043894.  
XX  
XX (KERJ ) FORSCHUNGSZENTRUM JUBILICH GMBH.  
XX  
XX Peters-Wendisch P, Eikmanns B, Salm H;  
XX  
XX WPI: 1999-245521/21.  
XX P-PSDB; AAW93971.  
XX  
XX Increasing microbial production of specific amino acids by increasing  
XX activity or expression of pyruvate carboxylase.  
XX  
XX Claim 16; Page 8-11; 18pp; German.  
XX  
XX This invention describes the isolation of a pyruvate carboxylase from  
XX Corynebacterium glutamicum which is used in a novel method for production  
XX of lysine, threonine, homoserine, glutamate and/or arginine, variously  
XX useful as feed additives, condiments, pharmaceuticals and intermediates  
XX for fine chemicals. Increasing pyruvate carboxylase activity increases  
XX the yield of microbial production of amino acids of the aspartate and/or  
XX glutamate families, e.g. about 50% more lysine, 40% more threonine and  
XX 150% more homoserine are secreted into the culture medium  
XX  
SQ Sequence 3728 BP; 813 A; 1059 C; 1035 G; 821 T; 0 U; 0 Other;  
  
Query Match 98.8%; Score 3576.4; DB 2; Length 3728;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 44 GCTTGAAGTCGTGACAGTTCAGGGGAGTGTGCGCGAACAATTGAGAGGAACAACAAAC 103  
Db 10 GCTTGAAGTCGTGACAGTTCAGGGGAGTGTGCGCGAACAATTGAGAGGAACAACAAAC 69  
  
QY 104 CGATGTTTGAATGGGGGAATCGGGGGTTACGATACTAGGACGAGTCACTGCTTATCACCC 163  
Db 70 CGATGTTTGAATGGGGGAATCGGGGGTTACGATACTAGGACGAGTCACTGCTTATCACCC 129  
  
QY 164 TTGGCGTCTCTGTTGAAGAGGATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 223  
Db 130 TTGGCGTCTCTGTTGAAGAGGATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 189  
  
QY 224 TTCCAGCATTCACAAAGATCTTGTAGCAACCGCGCGAATCGCGGTCGCTTTTCC 283  
Db 190 TTCCAGCATTCACAAAGATCTTGTAGCAACCGCGCGAATCGCGGTCGCTTTTCC 249  
  
QY 284 GTGCAGCACTCGAAACCGGTGACGCGACGCTAGCTATTTACCCCGGTGAAGATCGGGGAT 343  
Db 250 GTGCAGCACTCGAAACCGGTGACGCGACGCTAGCTATTTACCCCGGTGAAGATCGGGGAT 309

QY 344 CATTCCACCGCTCTTTTGGCTTCTGAAGCTGTCCGATTTGGTACCGAGGCTCACCACTCA 403  
Db 310 CATTCCACCGCTCTTTTGGCTTCTGAAGCTGTCCGATTTGGTACCGAGGCTCACCACTCA 369  
  
QY 404 AGCGGTACCTGGACATCATGATAAATTATCGGTGAGCTAAAAAGTTAAAGCAGATGCCA 463  
Db 370 AGCGGTACCTGGACATCATGATAAATTATCGGTGAGCTAAAAAGTTAAAGCAGATGCCA 429  
  
QY 464 TTTTACCGGGATACGGCTTCTCTGATAAATCCAGCTTCCCGGAGTGTGCGGAAA 523  
Db 430 TTTTACCGGGATACGGCTTCTCTGATAAATCCAGCTTCCCGGAGTGTGCGGAAA 489  
  
QY 524 ACGGCATTTACTTTTATTTGGCCCAACCCGAGAGTTCTTGATCTCACCGGTGATAAGTCTC 583  
Db 490 ACGGCATTTACTTTTATTTGGCCCAACCCGAGAGTTCTTGATCTCACCGGTGATAAGTCTC 549  
  
QY 584 CGCGGTAAACCGCGCGAGAGAGCTGCTCTGACAGTTTGGCGGANTCCACCCCGAGCA 643  
Db 550 CGCGGTAAACCGCGCGAGAGAGCTGCTCTGACAGTTTGGCGGANTCCACCCCGAGCA 609  
  
QY 644 AAAACATCGATGAGATCGTTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAG 703  
Db 610 AAAACATCGATGAGATCGTTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAG 669  
  
QY 704 CAGTTGCGGTGGTGGCGAGCGGTATGCTTTTGTGCTTACCTGATGAGCTTCGCA 763  
Db 670 CAGTTGCGGTGGTGGCGAGCGGTATGCTTTTGTGCTTACCTGATGAGCTTCGCA 729  
  
QY 764 AATTAGCAACAGAGCATCTGCTGAAGCTGAAGCGCTTTCGGCGATGCGCGGTATATG 823  
Db 730 AATTAGCAACAGAGCATCTGCTGAAGCTGAAGCGCTTTCGGCGATGCGCGGTATATG 789  
  
QY 824 TCGAAGCTGCTGATTAACCTCAGCATATTTGAAGTGCAGATCTCTTGGCGATCACA 883  
Db 790 TCGAAGCTGCTGATTAACCTCAGCATATTTGAAGTGCAGATCTCTTGGCGATCACA 849  
  
QY 884 GAGAAGTTGTACACCTTTTATGAACGTGACTGCTCAGTCAGCGCTGCTCACCAAAAGTTG 943  
Db 850 GAGAAGTTGTACACCTTTTATGAACGTGACTGCTCAGTCAGCGCTGCTCACCAAAAGTTG 909  
  
QY 944 TCGAATTTGCCCGCAGCAGCATTTTGAATCCAGAACTGCGTATCGCATTTGTGCGGATG 1003  
Db 910 TCGAATTTGCCCGCAGCAGCATTTTGAATCCAGAACTGCGTATCGCATTTGTGCGGATG 969  
  
QY 1004 CAGTAAAGTTCTGCGGTCTCATTGTTTACAGGCGCGGAAACCGTGAATCTTGTGTCG 1063  
Db 970 CAGTAAAGTTCTGCGGTCTCATTGTTTACAGGCGCGGAAACCGTGAATCTTGTGTCG 1029  
  
QY 1064 ATGAAAAGGGCAACACCGTCTTCAATCGAAATGAACCCACCTATCCAGGTTGAGCACCG 1123  
Db 1030 ATGAAAAGGGCAACACCGTCTTCAATCGAAATGAACCCACCTATCCAGGTTGAGCACCG 1089  
  
QY 1124 TGACTGAAGAGTCAACCGAGTGGACCTGCTGAGGCGCAGATCGCTTGGCTGCTGGTG 1183  
Db 1090 TGACTGAAGAGTCAACCGAGTGGACCTGCTGAGGCGCAGATCGCTTGGCTGCTGGTG 1149  
  
QY 1184 CAACCTTTGAAGGAATTTGGGTCTGACCCCAAGATTAAGATCAAGACCCACCGTGCAGCACTGC 1243  
Db 1150 CAACCTTTGAAGGAATTTGGGTCTGACCCCAAGATTAAGATCAAGACCCACCGTGCAGCACTGC 1209  
  
QY 1244 AGTCCGCTACCAACCGAGATCCAAACACCGCTTCCGCCAGATACCGGAACTATCA 1303  
Db 1210 AGTCCGCTACCAACCGAGATCCAAACACCGCTTCCGCCAGATACCGGAACTATCA 1269  
  
QY 1304 CCGGTACCGCTCACACGAGGAGCTGGCGTTCGCTTTGACGCTGACGCTCAGCTCGGTG 1363  
Db 1270 CCGGTACCGCTCACACGAGGAGCTGGCGTTCGCTTTGACGCTGACGCTCAGCTCGGTG 1329  
  
QY 1364 GCGAATTCACCGCATTCTTGACTTCCATGCTGTTGAAGTGAAGTCCGCTGCTGCGACT 1423  
Db 1330 GCGAATTCACCGCATTCTTGACTTCCATGCTGTTGAAGTGAAGTCCGCTGCTGCGACT 1389  
  
QY 1424 TTGAAACTGCTGTTGCTCGTGCACAGCGCGTGTGGCTGAGTTTACCGCTGTCTGCTGTTG 1483



181	ATTGGTACCGAAGGCTCACACAGCTCAAGCGGTACTACCTGGACATCGATGAATAATTATCCGGTGCA	240
Qy		
439	GCTAAAAAGTTAAAGCAGATGCCATTTTACC CGGATAGCGCTTCCTCTCTGTA AATGCC	498
Db		
241	GCTAAAAAGTTAAAGCAGATGCCATTTTACC CGGATAGCGCTTCCTCTCTGTA AATGCC	300
Qy		
499	CAGCTTCCCGCGAGTGTGCGGAAAACGGCATTTACTTTTATTGGCCCAACCCAGAGTT	558
Db		
301	CAGCTTCCCGCGAGTGTGCGGAAAACGGCATTTACTTTTATTGGCCCAACCCAGAGTT	360
Qy		
559	CTTGATCTCACCGGTGATTAAGTCTCCGCGGTAAACGCCGCGAAGAGCGTGTGCGCA	618
Db		
361	CTTGATCTCACCGGTGATTAAGTCTCCGCGGTAAACGCCGCGAAGAGCGTGTGCGCA	420
Qy		
619	GTTTTGGCGGAATCCACCCCGACGAAAAACATCGATGAGATGTTTAAAGCGCTGAAGGC	678
Db		
421	GTTTTGGCGGAATCCACCCCGACGAAAAACATCGATGAGATGTTTAAAGCGCTGAAGGC	480
Qy		
679	CAGACTTACCCCATCTTTTGTGAAGGCAGTTGCGGTGGTGGCGGACGCGGTATGCGTTTT	738
Db		
481	CAGACTTACCCCATCTTTTGTGAAGGCAGTTGCGGTGGTGGCGGACGCGGTATGCGTTTT	540
Qy		
739	GTGTGTTCACTGATGAGCTTCCGAAATTAGCAACGAAGCATCTCGTGAAGCTGAAGCG	798
Db		
541	GTGTGTTCACTGATGAGCTTCCGAAATTAGCAACGAAGCATCTCGTGAAGCTGAAGCG	600
Qy		
799	GCTTTTCGGCGATGCGCGGTATATGTGCAACGTGTGTGATTAAACCTTCAGCATATTGAA	858
Db		
601	GCTTTTCGGCGATGCGCGGTATATGTGCAACGTGTGTGATTAAACCTTCAGCATATTGAA	660
Qy		
859	GTGCAGATCTTTGGCGATCACATGCGAGAGTTGTACACCTTTATGAAGCTGACTGTCA	918
Db		
661	GTGCAGATCTTTGGCGATCACATGCGAGAGTTGTACACCTTTATGAAGCTGACTGTCA	720
Qy		
919	CTGCAGCGTCTGTCCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAA	978
Db		
721	CTGCAGCGTCTGTCCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAA	780
Qy		
979	CTGCGTGATGCATTTCTGCGGATGCAGTAAAGTCTCGCGCTCCATTGTTTACCAGGCG	1038
Db		
781	CTGCGTGATGCATTTCTGCGGATGCAGTAAAGTCTCGCGCTCCATTGTTTACCAGGCG	840
Qy		
1039	CGCGAAACCGTGGAAATCTTTGGTGCATGAAAGGGCAACACGCTTTTCATCGAAATGAAC	1098
Db		
841	CGCGAAACCGTGGAAATCTTTGGTGCATGAAAGGGCAACACGCTTTTCATCGAAATGAAC	900
Qy		
1099	CCAGCTATCCAGGTTGAGCACACCGTGACTGAAGAGTCAACGAGGTGACCTGCTGAAG	1158
Db		
901	CCAGCTATCCAGGTTGAGCACACCGTGACTGAAGAGTCAACGAGGTGACCTGCTGAAG	960
Qy		
1159	GGCGAGATGCGCTTGGCTGTGTGTGCMACTTTGAAGGAATTTGGTCTGACCCAGATTAAG	1218
Db		
961	GGCGAGATGCGCTTGGCTGTGTGTGCMACTTTGAAGGAATTTGGTCTGACCCAGATTAAG	1020
Qy		
1219	ATCAAGACCCACGGTCAGCACCTGCAGTGGCGCATCCACGGAAGATCCAAACCAACGGC	1278
Db		
1021	ATCAAGACCCACGGTCAGCACCTGCAGTGGCGCATCCACGGAAGATCCAAACCAACGGC	1080
Qy		
1279	TTCCGCCAGATACCGAACTATATCACCGCGTACCGTCTACGAGCGGAGCTTGCGCTTGGT	1338
Db		
1081	TTCCGCCAGATACCGAACTATATCACCGCGTACCGTCTACGAGCGGAGCTTGCGCTTGGT	1140
Qy		
1339	CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTTGACTTCCATGCTGGTG	1398
Db		
1141	CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTTGACTTCCATGCTGGTG	1200
Qy		
1399	AAAAATGACTCGCGTGGTTCGGACTTTGAACTGTGTTGCTCGTGCAAGCGCGGTTG	1458
Db		
1201	AAAAATGACTCGCGTGGTTCGGACTTTGAACTGTGTTGCTCGTGCAAGCGCGGTTG	1260
Qy		
1459	GCTGAGTTTCAACCGTGTCTGGTGTGTCGAAACCAACATTTGGTTTCTTGGTGGCTGCTGCGG	1518

Db 1261 GCTGAGTTACCGTGTCTGGTGTGCAACCAACATGTTGTTTCTTGGTGGTCTGCTGGG 1320  
Qy 1519 GAAGAGGACTTCACTTCCAAAGCGATCGCCACCGGATTCATTGGCGGATCACCCGACATC 1578  
Db 1321 GAAGAGGACTTCACTTCCAAAGCGATCGCCACCGGATTCATTGGCGGATCACCCGACATC 1380  
Qy 1579 CTTGAGGCTCACCTGCTGATGATGAGCAGGACGATCCTGGATTTACTTGGCAGATGTC 1638  
Db 1381 CTTGAGGCTCACCTGCTGATGATGAGCAGGACGATCCTGGATTTACTTGGCAGATGTC 1440  
Qy 1639 ACCGTGAACAAGCCTCATGGTGTGGTCCAAAGGATGTTGAGGCTTCCTATCGATAAGCTG 1698  
Db 1441 ACCGTGAACAAGCCTCATGGTGTGGTCCAAAGGATGTTGAGGCTTCCTATCGATAAGCTG 1500  
Qy 1699 CTTAATCATCAAGGATCTGCACTGCGACCGGTTCCCGTGACCGGCTGAAGAGCTTGGC 1758  
Db 1501 CTTAATCATCAAGGATCTGCACTGCGACCGGTTCCCGTGACCGGCTGAAGAGCTTGGC 1560  
Qy 1759 CAGCGCGGTTTGTCTGATCTCGTGAGCAGGACGACCTGGGAGCTTACTGATACCACTC 1818  
Db 1561 CAGCGCGGTTTGTCTGATCTCGTGAGCAGGACGACCTGGGAGCTTACTGATACCACTC 1620  
Qy 1819 TTTCCGCGATGACACACAGCTTTTGTGTCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1878  
Db 1621 TTTCCGCGATGACACACAGCTTTTGTGTCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1680  
Qy 1879 GCGGACAGGCGCTGCGCAAGCTGACTCTCGTGAGCAGGACGACCTGGGAGCTTACTGATACCACTC 1938  
Db 1681 GCGGACAGGCGCTGCGCAAGCTGACTCTCGTGAGCTTTTGTCTGAGCTTTTGTCTGGGAGGCTTGGGCGGC 1740  
Qy 1939 GCGACTACGATGTGGCGATGCTTTTCTCTTTGAGGATTCGTTGGGACAGGCTCGACGAG 1998  
Db 1741 GCGACTACGATGTGGCGATGCTTTTCTCTTTGAGGATTCGTTGGGACAGGCTCGACGAG 1800  
Qy 1999 CTGGCGAGGCGATCGCGAATCTAAACATTCAGATGCTGCTTCGCGGCGGCAACACCGTG 2058  
Db 1801 CTGGCGAGGCGATCGCGAATCTAAACATTCAGATGCTGCTTCGCGGCGGCAACACCGTG 1860  
Qy 2059 GGATACACCCCGTACCAGACTCCGTCTGCGCGGCTTTGTTAAGGAAGCTGCCAGCTCC 2118  
Db 1861 GGATACACCCCGTACCAGACTCCGTCTGCGCGGCTTTGTTAAGGAAGCTGCCAGCTCC 1920  
Qy 2119 GCGTGGACATCTTCGGATCTTCGACGCGCTTAACGAGCTCCCGAGTGGTCCAGCA 2178  
Db 1921 GCGTGGACATCTTCGGATCTTCGACGCGCTTAACGAGCTCCCGAGTGGTCCAGCA 1980  
Qy 2179 ATCGACGAGTCTCTGAGACCAACACCGCGGTAGCGAGTGGCTATGGCTTATTCTGGT 2238  
Db 1981 ATCGACGAGTCTCTGAGACCAACACCGCGGTAGCGAGTGGCTATGGCTTATTCTGGT 2040  
Qy 2239 GATCTCTGTATCCAAATGAAAAGCTTACACCTTGGATTTACTCTAAAGATGCGAG 2298  
Db 2041 GATCTCTGTATCCAAATGAAAAGCTTACACCTTGGATTTACTCTAAAGATGCGAG 2100  
Qy 2299 GAGATCGTCAAGTCTGGCGCTCACATCTTTGGCCATTAAGGATATGGCTGGTCTGCTCGC 2358  
Db 2101 GAGATCGTCAAGTCTGGCGCTCACATCTTTGGCCATTAAGGATATGGCTGGTCTGCTCGC 2160  
Qy 2359 CAGCTGCGGTAAACCAAGCTGTGTCACCGCACTGCGCGGTGAATTCGATCTGCGAGTGCA 2418  
Db 2161 CAGCTGCGGTAAACCAAGCTGTGTCACCGCACTGCGCGGTGAATTCGATCTGCGAGTGCA 2220  
Qy 2419 GTGCACACCCAGCAGCTGCGGTGGCCAGCTGGCAACCTACTTTGCTGCACTCAAGCT 2478  
Db 2221 GTGCACACCCAGCAGCTGCGGTGGCCAGCTGGCAACCTACTTTGCTGCACTCAAGCT 2280  
Qy 2479 GTGTCAGATGCTGTGAGCGGCTTCGCAACCACTGTCTGGCAACCACTTCGAGCCATCC 2538  
Db 2281 GTGTCAGATGCTGTGAGCGGCTTCGCAACCACTGTCTGGCAACCACTTCGAGCCATCC 2340  
Qy 2539 CTGCTCGGCAATGTTGCTGCTCATTCGCGCACACCCGTCGCGATACCGGTTTGGAGCTCGAG 2598  
Db 2341 CTGCTCGGCAATGTTGCTGCTCATTCGCGCACACCCGTCGCGATACCGGTTTGGAGCTCGAG 2400

RESULT 8

ABK52832

ID ABK52832 standard; DNA; 3474 BP.

Qy 2599 GCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGGCGGACTGTACTGCAATTTGAG 2658  
Db 2401 GCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGGCGGACTGTACTGCAATTTGAG 2460  
Qy 2659 TCTGGAAACCCAGGCCCACCGTCTGCTCTACCGCCAGCAATCCAGGCGGACAGTTG 2718  
Db 2461 TCTGGAAACCCAGGCCCACCGTCTGCTCTACCGCCAGCAATCCAGGCGGACAGTTG 2520  
Qy 2719 TCCAACTCGTGGTGCACAGGCCACCGCACTGGGCTTTGGCGATCGTTTGGATCTCATCGAA 2778  
Db 2521 TCCAACTCGTGGTGCACAGGCCACCGCACTGGGCTTTGGCGATCGTTTGGATCTCATCGAA 2580  
Qy 2779 GACAACTAGCGAGCGTTTAAATGAGATCTGGGACGCCCAACCAAGGTCAACCCATCTCTCC 2838  
Db 2581 GACAACTAGCGAGCGTTTAAATGAGATCTGGGACGCCCAACCAAGGTCAACCCATCTCTCC 2640  
Qy 2839 AAGTTGTTGGGACCTCGCACTCCACTCTGTTGGTGGCGGTGTGGATCCAGCAGACTTT 2898  
Db 2641 AAGTTGTTGGGACCTCGCACTCCACTCTGTTGGTGGCGGTGTGGATCCAGCAGACTTT 2700  
Qy 2899 GCTGCGGATCCCAAAAGTACGACATCCAGACTCTGTCTATCGCTTCTGCGCGGCGAG 2958  
Db 2701 GCTGCGGATCCCAAAAGTACGACATCCAGACTCTGTCTATCGCTTCTGCGCGGCGAG 2760  
Qy 2959 CTTGGTAAACCTCCAGTGGCTGGCCAGAGCCACTGCGCACCCCGGCACTGGAAAGGCCGC 3018  
Db 2761 CTTGGTAAACCTCCAGTGGCTGGCCAGAGCCACTGCGCACCCCGGCACTGGAAAGGCCGC 2820  
Qy 3019 TCCGAAGGCAAGGCACTCTGACGGAAGTTCTTGAGGAAGCAGCGCCACCTCGACGCT 3078  
Db 2821 TCCGAAGGCAAGGCACTCTGACGGAAGTTCTTGAGGAAGCAGCGCCACCTCGACGCT 2880  
Qy 3079 GATGATTCCAAAGGAAGTGGCAATAGCTCAACCGCTCTGTTCCTCCGAAGCCAAACCGAA 3138  
Db 2881 GATGATTCCAAAGGAAGTGGCAATAGCTCAACCGCTCTGTTCCTCCGAAGCCAAACCGAA 2940  
Qy 3139 GAGTTCTCTGAGCAACCGTCCGCTTTGGCAACACCTCTGCGCTGATGATCGTGAATTC 3198  
Db 2941 GAGTTCTCTGAGCAACCGTCCGCTTTGGCAACACCTCTGCGCTGATGATCGTGAATTC 3000  
Qy 3199 TTTCTAGCGCTGTGTCGAAGCGCGGAGACTTTTGTATCCGCTGCGAGATGTGCGACCCCA 3258  
Db 3001 TTTCTAGCGCTGTGTCGAAGCGCGGAGACTTTTGTATCCGCTGCGAGATGTGCGACCCCA 3060  
Qy 3259 CTGCTTTGCTGCTGATGATGATCTCTGAGCCAGCAGATAAGGATATGCGCAATGTTGTG 3318  
Db 3061 CTGCTTTGCTGCTGATGATGATCTCTGAGCCAGCAGATAAGGATATGCGCAATGTTGTG 3120  
Qy 3319 GCTCAAGTCAACCGCCAGATCCGCCAATGCGTGTGCGTGAACCGCTCCGTTGAGTCTGTC 3378  
Db 3121 GCTCAAGTCAACCGCCAGATCCGCCAATGCGTGTGCGTGAACCGCTCCGTTGAGTCTGTC 3180  
Qy 3379 ACCGCAACCGCAGAAAAGCAGATTCTCTCAACAAAGGGCCATGTTGCTGCAACCATTCGCT 3438  
Db 3181 ACCGCAACCGCAGAAAAGCAGATTCTCTCAACAAAGGGCCATGTTGCTGCAACCATTCGCT 3240  
Qy 3439 GGTGTTGTACCGTGTGATGTTGCTGAGGTGATGAGGTCAAGGCTGAGATGCGATCGCA 3498  
Db 3241 GGTGTTGTACCGTGTGATGTTGCTGAGGTGATGAGGTCAAGGCTGAGATGCGATCGCA 3300  
Qy 3499 ATCATCGAGGCTATGAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATTCGAT 3558  
Db 3301 ATCATCGAGGCTATGAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATTCGAT 3360  
Qy 3559 CGGTTGTGTTCTCTGCTGCAACGAAGGTGGAAGGTGGGAGCTTGATCGTCTGCTTTC 3618  
Db 3361 CGGTTGTGTTCTCTGCTGCAACGAAGGTGGAAGGTGGGAGCTTGATCGTCTGCTTTC 3420

XX AC ABK52832;  
 XX DT 27-AUG-2002 (first entry)  
 XX DE Corynebacterium glutamicum feedback-resistant pyruvate carboxylase gene.  
 XX KW Feedback-resistant; pyruvate carboxylase enzyme; gene; ds;  
 XX KW aspartic acid feedback inhibition resistant.  
 XX OS Corynebacterium glutamicum.  
 XX FH Key Location/Qualifiers  
 XX FT CDS 1. 3474  
 FT /tag= a  
 FT /product= "Corynebacterium glutamicum feedback-resistant  
 FT pyruvate carboxylase enzyme"  
 FT /transl\_except= (pos:1..3; aa:Met)  
 XX WO200231158-A2.  
 XX 18-APR-2002.  
 XX 12-OCT-2001; 2001WO-US031893.  
 XX 13-OCT-2000; 2000US-0239913P.  
 XX (ARCH ) ARCHER-DANIELS MIDLAND CO.  
 XX Hanke PD;  
 XX WPI; 2002-463267/49.  
 XX P-PSDB; AAU98050.  
 XX Novel mutated, feedback resistant pyruvate carboxylase enzyme  
 XX polypeptide, useful for producing amino acids e.g. L-lysine, L-threonine,  
 XX L-glycine, L-glutamic acid, L-proline and L-methionine and L-isoleucine.  
 XX Claim 3; Fig 1; 42pp; English.  
 XX The present invention relates to a new mutated, feedback-resistant  
 XX pyruvate carboxylase enzyme. The invention is useful for producing an  
 XX amino acid (e.g. L-lys, L-thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro), by  
 XX culturing a host cell in a suitable media and separating the amino acid  
 XX from the medium. The vector of the invention is useful for replacement of  
 XX a wild-type pyruvate carboxylase gene, with a feedback resistant pyruvate  
 XX carboxylase gene, in a Corynebacterium glutamicum, by replacing a genomic  
 XX copy of the wild-type pyruvate carboxylase gene with a selectable marker  
 XX gene through homologous recombination to form a first recombinant  
 XX strain, and replacing the selectable marker gene in the first recombinant  
 XX strain, with feedback resistant pyruvate carboxylase gene through  
 XX homologous recombination to form a second recombinant strain, where the  
 XX homologous recombination in the above steps, occurs between the host cell  
 XX and the vector. The feedback-resistant pyruvate carboxylase enzyme is  
 XX resistant to feedback inhibition from aspartic acid. The present nucleic  
 XX acid sequence encodes the feedback-resistant pyruvate carboxylase enzyme  
 XX of the invention  
 XX SQ Sequence 3474 BP; 749 A; 1013 C; 950 G; 762 T; 0 U; 0 Other;  
 Query Match 93.9%; Score 3398.8; DB 6; Length 3474;  
 Best Local Similarity 98.6%; Pred. No. 0;  
 Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
 QY 148 GTGACTGTATCACCTTGGCGGTCTCTTGTGAAGGAATAATTACTCTAGTGTGACT 207  
 DB 1 GTGACTGTATCACCTTGGCGGTCTCTTGTGAAGGAATAATTACTCTAGTGTGACT 60  
 QY 208 CACACATCTTCAACGCTTCCAGCATTTCAAAAGATCTTTGGTAGCAAAACCGCGCGAAATC 267  
 DB 61 CACACATCTTCAACGCTTCCAGCATTTCAAAAGATCTTTGGTAGCAAAACCGCGCGAAATC 120  
 QY 268 GCGGTCCGTGCTTCCGTGAGCACTCGAAACCGGTGACCCACGCGTAGCTATTACCCC 327

DB 121 GCGGTCCGTGCTTCCGTGAGCACTCGAAACCGGTGACGCCAGGTAGCTATTATACCC 180  
 QY 328 CGTGAAGATCGGGATCATTTCCACCGCTCTTTTGTCTTGAAGCTGTCGCGATTGGTACC 387  
 DB 181 CGTGAAGATCGGGATCATTTCCACCGCTCTTTTGTCTTGAAGCTGTCGCGATTGGTACT 240  
 QY 388 GAAGGCTCACAGTCAAGGGGTACCTCGACATCGATGAATTAATTCGCTGAGCTTAAAAA 447  
 DB 241 GAAGGCTCACAGTCAAGGGGTACCTCGACATCGATGAATTAATTCGCTGAGCTTAAAAA 300  
 QY 448 GTTAAAGCAGATGCCATTATACCGGGATAGGCTTCTGCTGTAATATGCCAGCTTGCC 507  
 DB 301 GTTAAAGCAGATGCTATTATACCGGGATATGGCTTCTGCTGTAATATGCCAGCTTGCC 360  
 QY 508 CGCAGTGTGCGAAAAACGGCATTTATTTATTTGGCCCAACCCACAGAGTTCCTGATCTC 567  
 DB 361 CGCAGTGTGCGAAAAACGGCATTTATTTATTTGGCCCAACCCACAGAGTTCCTGATCTC 420  
 QY 568 ACCGTTGATAGTCTCGCGGTTAACCGCGGAAGAGGCTGCTGCGAGTTTTCGCG 627  
 DB 421 ACCGTTGATAGTCTCGCGGTTAACCGCGGAAGAGGCTGCTGCGAGTTTTCGCG 480  
 QY 628 GAATCCACCCGAGCAAAAAACATCGATGAGATCGTTAAAAAGCGCTGAAGCCAGACTTAC 687  
 DB 481 GAATCCACCCGAGCAAAAAACATCGATGAGATCGTTAAAAAGCGCTGAAGCCAGACTTAC 540  
 QY 688 CCCATCTTTGTGAAGCAGTTGCGGTTGCGGACGCGGTATGCGCTTTTGTGCTTCA 747  
 DB 541 CCCATCTTTGTGAAGCAGTTGCGGTTGCGGACGCGGTATGCGCTTTTGTGCTTCA 600  
 QY 748 CCTGATGAGCTTCGCAAAATTAGCAACAGAAGCATCTCTGTAAGCTGAAGCGGCTTCGCG 807  
 DB 601 CCTGATGAGCTTCGCAAAATTAGCAACAGAAGCATCTCTGTAAGCTGAAGCGGCTTCGCG 660  
 QY 808 GATGCGCGGTATATGCAAGCTGCTGTGTAATTAACCTCAGCATATTTGAGTCCAGATC 867  
 DB 661 GACGGTTCGGTATATGCAAGCTGCTGTGTAATTAACCCCGAGCATTTGAAGTGCAGATC 720  
 QY 868 CTTGCGCATCACACTGGAGAAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 927  
 DB 721 CTTGCGCATCGACTGGAGAAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780  
 QY 928 GGTCAACAAAAGTTGTGCAAAATTGCGGACGACAGCATTTGGATCCAGAACTCGGTGAT 987  
 DB 781 GGTCAACAAAAGTTGTGCAAAATTGCGGACGACAGCATTTGGATCCAGAACTCGGTGAT 840  
 QY 988 CGCATTTGTGCGGATGCAAGTAAAGTTCTGCGCTCCATTGGTTACCAAGGCGCGGAACC 1047  
 DB 841 CGCATTTGTGCGGATGCAAGTAAAGTTCTGCGCTCCATTGGTTACCAAGGCGCGGAACC 900  
 QY 1048 GTTGAATTTCTTGGTGCATGAAAGGCGCAACACAGTCTTTCATCGAAATGAACCCAGTATC 1107  
 DB 901 GTTGAATTTCTTGGTGCATGAAAGGCGCAACACAGTCTTTCATCGAAATGAACCCAGTATC 960  
 QY 1108 CAGGTTGAGCACACCGTGACTGAAAGAGTCAACGAGGTGGACCTGGTGAAGGCGCGAGATG 1167  
 DB 961 CAGGTTGAGCACACCGTGACTGAAAGAGTCAACGAGGTGGACCTGGTGAAGGCGCGAGATG 1020  
 QY 1168 CGCTTGGTGTGTGTGCAACCTTGAAGGAATTGGTCTGACCCAGATTAAGATCAAGAAC 1227  
 DB 1021 CGCTTGGTGTGTGTGCAACCTTGAAGGAATTGGTCTGACCCAGATTAAGATCAAGAAC 1080  
 QY 1228 CACGGTGACGACTGCACTGCGGATCACACGGAAGATCCAAACACCGGCTTCGCGCCA 1287  
 DB 1081 CACGGTGACGACTGCACTGCGGATCACACGGAAGATCCAAACACCGGCTTCGCGCCA 1140  
 QY 1288 GATACCGGAATATCACCGGTACCGCTCAACAGCGGAGCTGGCGTTTCGTTGACGGT 1347  
 DB 1141 GATACCGGAATATCACCGGTACCGCTCAACAGCGGAGCTGGCGTTTCGTTGACGGT 1200  
 QY 1348 GCAGCTCAGCTCGGTGGGAATCACCGCACATTTGACTCCATGCTGCTGTAATATGACC 1407







Db 1325 TGACCCAGGACAAGATACACCAACCGGTGCGCCCTGCAGTGCAGCATCACCGAGG 1384  
Qy 1265 ATCCAAACAAAGCTTTCGCCCCAGATACCGAACTATCACCGCGTACCGCTCACAGCG 1324  
Db 1385 ACCCGTCCAAACAACTTCGCGCCCGACACCGGTGTGATCACCGCTACCGCTCCCGGGT 1444  
Qy 1325 GAGCTGGCGTTTCGTTGACGCTGACGCTGAGCTGAGTCCGTTGGGGAATACACCGACACTTTG 1384  
Db 1445 GTGCGGGTGTGGTCTCGACGCGCGAGCCAGCTCGGCGGAGATCACCGACATTTTCG 1504  
Qy 1385 ACTCCATGCTGTGAAATGACTCGCGGTGTTCCGACTTTGAACTGCTGTGTCGTG 1444  
Db 1505 ATTCCATGCTGTCAAGATGACTTCGCGGGTTCGATTTTCGAGACCGCGGTTCGCGAG 1564  
Qy 1445 CACAGCGCGTTGCTGATGATTTCAACCGGTTCGAGTTCGAGTTCGAACTGCTGTTTCGTC 1504  
Db 1565 CCGAGCGCGCTGCGGAGTTTCAACGCTTCGCGGTGCGCACCAACATCGGCTTCCTCG 1624  
Qy 1505 GTGCGTTGCTGCGGAAGAGACTTCACTTCCAAAGCGATCGCACCGGATTCATTCGCG 1564  
Db 1625 GTGCGTGTGCTGCGAGGAGACTTCAACAGAGCGGATCGACACCGGCTTCATCGGCT 1684  
Qy 1565 ATCACCGCACCTCTCAGGCTCAACCTGCTGATGATGAGCAGGAGCGATTCCTGGAT 1624  
Db 1685 CCGACAGCACTGCTCCAGGCCCCACCGCGCAGATGAGCAGGGCGGATCCTGGAAT 1744  
Qy 1625 ACTGGCAGATGCTACCGGTGACAGCCCTCAAGTGTGGGTCCAAAGGATGTCAGGTC 1684  
Db 1745 ACTGCGGATGTACCGGTGAACAAACCCAGCGGTGAACGCC---CGAGACAGCCGCTC 1801  
Qy 1685 CTATCGATAGCTGCCTTAACATCAAGGATCTGCCACTGCGACGCGTTCCGCTGACCGCC 1744  
Db 1802 CGATAGAGAGCTGCCCGAGGTGGAGAACATCCCGCTGCCACGCGCTCCCGACCGCC 1861  
Qy 1745 TGAAGAGCTTGGCCCCAGCGGGTTGTCGAGTCTCGTGGAGCAGGACGACTGGCAG 1804  
Db 1862 TGAAGAGCTCGGCGCGAGGTTTCGCCCCGATCTGCGAAACAGGATGCCCTGGCGG 1921  
Qy 1805 TTACTGATACCACTTTCGCGATGACACACAGCTTTGTTGCGACCGCGAGTCCGCTCAT 1864  
Db 1922 TCACGACACCACTTTCGCGATGCCACCACTCCCTCTGSCACCCCGCTGCGCTCT 1981  
Qy 1865 TGCAGTGAAGCTTGGCGAGAGCGGTGCGAAAGTGAATCTCGAGCTTTTGTCCGTGG 1924  
Db 1982 TCGCGTGAACCCCGCGCGCGCGCGTGCAGAACTCAACCCCGAGCTGCTGTCGCTGG 2041  
Qy 1925 AGGCTTGGCGCGCGGACTTACGATGTGGGATGCTTCTCTTTCGAGATCCGTTGGG 1984  
Db 2042 AGGCTTGGCGCGGTGCCACTTACGACGTGGCCATGGCTTCTTTCGAGGATCCGTGGG 2101  
Qy 1985 ACAGGCTCGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCG 2044  
Db 2102 CACGCTGATGAGCTGCGTGAAGCGGATGCCGAATGTGAACATTCAGATGCTGCTGCGTG 2161  
Qy 2045 GCGCAACACCGTGGGATACACCGGTACCGAGCTCGCTGCGCGCGGTTGTTGAAG 2104  
Db 2162 GTGCGAAACCGTCCGGTACACCGGTACCGGTTCGGTGTGCGCGGTTTGTGCGAG 2221  
Qy 2105 AAGTCCAGCTTCGCGGTGGAATCTTCGCTTCGCTTCGAGCGGCTTAACGAGCTTCCTCC 2164  
Db 2222 AGGCGCCCAAGTCCCGGTGTGGAATCTTCGCTTCGAGCGGCTCAACGACATCTCC 2281  
Qy 2165 AGATGGTCCAGCAATCGAGCAGTCTCGAGACCAACACCGCGGTAGCGAGGTGGCTA 2224  
Db 2282 AGATGCGCCCGGCATCGAGCGGCTCTGGAGACCGGACCAAGTGTGCGAGGTGCGCA 2341  
Qy 2225 TGGCTTATTTCTGATCTCTCTGATTCACAAATGAAAGCTTACACCTTGATTAATACC 2284  
Db 2342 TGGCGTACTCCGCTGACCTGTCCAAATCCGCGGAGAGAGCTTACACCTTGAGTACTACC 2401  
Qy 2285 TAAAGTGCAGAGGAGTCTGAGTCTGGCGCTCAGATCTTGGCCATTAAGGATAGG 2344  
Db 2402 TGAACCTGCGGAGAGATCGTCGACTCCGGGTGCAACACATCTGGCCATCAGGACATGG 2461

Qy 2345 CTGCTCTGCTTGGCCAGCTGGGTAACCAAGCTGTGTACCGCACTGCGCCCTGATTCG 2404  
Db 2462 CCGGCTGTGCGCGCGCGCGGCCCAAACTGGTCAACCGCTGCGCGGTGATTCG 2521  
Qy 2405 ATCTGCAGTGCAGCTGCAACCCAGCAGACTGCGGGTGGCCAGCTGGCAACTACTTTG 2464  
Db 2522 ACCTGCGCGTGCATGTCCACACCCAGCAGACCGCGCGGCTGAGCTGGCAGCTACTCTCG 2581  
Qy 2465 CTCAGCTCAAGCTGTGTGAGATGCTGTGAGGGTCTTCCGACCACTGTCTGGCAGCA 2524  
Db 2582 CCGCGCCCAACCGCGCGCGGCGGATCGCTGACCGCGCTCCGACCCCTGTCCGGTACCA 2641  
Qy 2525 CTTCCAGCAGCTCCCTGTCTGTCATTTGCTGTCATTTCGCGCACACCGTTCGCGATACG 2584  
Db 2642 CTTCCAGCAGCTGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2701  
Qy 2585 GTTTGAGCTTCGAGGCTGTTCGAGCTCGAGCCGCTACTTGGGAAGCAGTGCAGGCTGT 2644  
Db 2702 GCTTCACTGAGGCGCTCTCCGACTGGAACCGTACTTGGGAGGGGTTCGCGGACTGT 2761  
Qy 2645 ACCTGCGATTTGAGTCTGGAACCCAGGCGCCAACCGGTGCGCTTACCGCGCACGAATCC 2704  
Db 2762 ACCTGCGTTTGAATCCGCGCACCCCGCGCGCCGACCGGACGCGTTTACCGCCACGATCC 2821  
Qy 2705 CAGCGGACAGTTGTCACCTGCGTGCAGCGCACACCGCACTGGGCTTGGCGATCGTT 2764  
Db 2822 CCGCGGTGAGCTGTCTCACTGCTGCTGCGCGCGCTTGCATGCGGTCTGCGCGCGCT 2881  
Qy 2765 TCGAATCTATCGAAGCAAACTAGCAGCGGTTAATGAGATGTGGGACGCGCCCAACAAAG 2824  
Db 2882 TCGAGCTATCGAGGACTACTAGCGCGCGCTCAACGAGATGCTGGGTCTGCGCACAAAG 2941  
Qy 2825 TCACCCATCTCCAAAGTTGTTGGGAGCTCGCACTCCACCTGCTGTTGGTGGGGTGTGG 2884  
Db 2942 TCACCCGCTCTCAAGTTGTCGAGTCTGCACTGCACTGCACTGCTGCTGCTGCTGCTG 3001  
Qy 2885 ATCCAGCAGACTTTGCTGCGGATCCACAAAGTACGACATCCAGACTCTGTCTATCGCGT 2944  
Db 3002 GCGCGGAGGATTCGCGCGCGATCCGAGAGTACGACATCCCGATTCGCTCATCGCT 3061  
Qy 2945 TCTGCGCGGAGCTTGGTAACTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCCGG 3004  
Db 3062 TCTTCCGCGGAGCTGGGTACCCCTCCCGGTGGCTGGCCAGACCGCTGCGCACCCCGG 3121  
Qy 3005 CACTGGAAGCCCGCTCCGAGGCAAGGCACTCTGACGGAAGTTCTTGAGGAGAGAGCAG 3064  
Db 3122 CACTCGAGGCTCGCTCCAGGGTAAGGCGCGCTGCGGAGATCCCGCGCGAGAGCAG 3181  
Qy 3065 CGCACTCGAGCTGATGATTCAGAGAGCTGCGAATAGCTTCAACCGCTGCTGCTTCC 3124  
Db 3182 CCACCTTGGATTCGATGATTCGCGGAGCTGCGGCGACCTCAACCGCTGCTGCTTCC 3241  
Qy 3125 CGAAGCAACCGAAGAGTTCTCTCGAGCACCGTTCGCGCTTCGGAACACCTCTGCGCTGG 3184  
Db 3242 CGAAGCGACCGAGGAGTTCTTTCGAGCACCGTTCGCGCTTCGGAACACCTCTGCGCTGG 3301  
Qy 3185 ATGATGATGATTTCTTCTAGCGCTGTCGAGGCGCGGAGACTTTGATTCGCTGCGCAG 3244  
Db 3302 ATGACGCGAGTTCTTCTAGCGCTTGAAGAGGAGCTGAGGAGCTGATCCGACTGACCG 3361  
Qy 3245 ATGTGCGCACCCCACTGCTTGTTCGCTGATTCGAGCTCTGAGCCAGACGATAAGGGTA 3304  
Db 3362 GTGTGTCACCCCGATGTTGTTGCTGCTGATGCGGTGTCGGAACCGGATGACAAAGGCA 3421  
Qy 3305 TGGCGAATTTGTTGGCCAACTGCAAGGCGCAGATCCGCGCAATGCGGTGCTGCTGAGCGCT 3364  
Db 3422 TGGCGCAACGTGTGTGTCAACGTCACGCGCAGATCCGCGCGATCAAGGTTGCGCGACCGTT 3481  
Qy 3365 CCGTGTGCTGTGTCAGCGCAACCGCAAGAAAGGAGATTCCTCCAAACAGGGGCAATGTTG 3424  
Db 3482 CCGTGTGAGTCCGTCAACCGCACCCCGGAGAGGCGCGATGCCCAACAGGGGCAATGTCG 3541

QY 3425 CTGACCAATTCGCTGTTGTCACCGTACCTTCTGTAAGGTGATGAGGTCAAGGCTG 3484  
Db 3542 CCGCACCAATTCGCGGTGTGTGTCACCGTACCGTCCGAGGGTGTATGAGATCAAGGCTG 3601  
QY 3485 GAGATGCAATTCGCAATCATCGAGGCTATGAGATGCAAGCAAACTACTGTTCTGTG 3544  
Db 3602 GCGACGCGTGGCCATCATTTGAGGCCATGAGATGAGGCCACCATCACCGGCGCTGTG 3661  
QY 3545 ACGGCAAAATCGATTCGCTGTTGTTCTGCTGCAACGAGGTGGAAGTGGGACTTGA 3604  
Db 3662 ACGGTGTATCGACCGCGTGTGTGTCGCGCCGCCACCAAGGTGCGAGGCGGCGACTCA 3721  
QY 3605 TCGTGTGCTGTTCTTA 3620  
Db 3722 TCGTGTGCTGTTCTTA 3737

RESULT 10

ACA29831  
ID ACA29831 standard; DNA; 3423 BP.  
AC ACA29831;  
DT 19-JUN-2003 (first entry)  
DE Prokaryotic essential gene #11488.  
KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW drug design; gene.  
OS Corynebacterium diphtheriae.  
FN WO200277183-A2.  
PD 03-OCT-2002.  
FF 21-MAR-2002; 2002WO-US009107.  
XX 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX (ELIT-) ELITRA PHARM INC.  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
DR P-FSDB; ABU25961.

PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 14; SEQ ID NO 17701; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 3423 BP; 773 A; 1058 C; 912 G; 680 T; 0 U; 0 Other;

Query Match 53.2%; Score 1927.4; DB 7; Length 3423;  
Best Local Similarity 73.4%; Pred. No. 0;  
Matches 2494; Conservative 0; Mismatches 896; Indels 9; Gaps 2;  
QY 213 ATCTTCACGCTTCCAGCATTCACAAAGATCTTGTAGCAACCGCGGCGAAATTCGGGT 272  
Db 27 AACCTCTACGGTAATCCATTTGTCAAAGATCTCTGTTGCTAACCGTGGCGAGATCGCGT 86  
QY 273 CCGTGTCTTCCGTGCGAGCATTCGAAACCGGTGCGACCGTAGCTATTTACCCCGGTGA 332  
Db 87 GCGAGCTTTTCGTGCGAGCTTTGAAACCGGTGCTGCCACGCTAGCTCTACCCGGAACGA 146  
QY 333 AGATCGGGATCATTCACCGCTCTTTTCTTCTGAGAGCTGTCGCGATTCGTCGGAAGG 392  
Db 147 AGACCGTAATTCATTCCACCGATCGTTTGGCTCTGAAAGCTGTCTCTATTGTTGAGGCGG 206  
QY 393 CTCACGCTCAAGCGTACCTGCGATCGATGAAATTCGCTGCGAGCTTCTTGATCTCACGG 452  
Db 207 ATCTGCGTCAAGGCTACCTCGATATCGACGAATCATCGTCTGCTGTAACAAACGGG 266  
QY 453 AGCAGATGCCATTTACCGGGATACGGTCTCTCTGTAATGCGCGAGTTCGCGCGA 512  
Db 267 TGCCGACGCAATTTACCCAGGCTACGGCTCTCTTCCGAAACCGCCAGCTTGTCTGTA 326  
QY 513 GTGTGCGGAAACCGCATTTACTTTTATTTGGCCCAACCCCGAGAGTTCTTGATCTCACGG 572  
Db 327 GTGCGTCAAAACGGCATACCTTTATCGTCTCCACCTCCATCAGTGTGGAGCTTACCG 386  
QY 573 TGATTAAGTCTCGCGGTAACCGCGCGGAGAGGCTGTCTGCGAGTTTGGCGGATC 632  
Db 387 TGATAAAGCAGCTGCTGTATCCGCGACGCGGAGGCTTGGCCAGCTGACTGAGAC 446  
QY 633 CACCCCGAGCAAAACATCGATGAGATCGTTTAAAGCGCTGAAGGCCAGACTTACCCCAT 692  
Db 447 AGAAGCAACAGCAGTCCAAAGAGCTGCGAGAGCTGACCAAGGGGCGAGACTACCCGCT 506  
QY 693 CTTTGTGAAGGAGTTCGGGTGTGGCGGACGCGGTATGCGTTTGTGTTTGTTCACCTGA 752  
Db 507 ATTCGTAAAGCGGTTCGGGTGTGGCGGCGCATGCGTTTGTGCGAGCGCCCGCA 566  
QY 753 TGAGTCTGCGAAATFAGCAACAGAGCATCTCGTGAAGCTGAAGCGGCTTTCGCGGATGG 812  
Db 567 AAACCTAGAAAGCTAGCAGCAGAGCTCCCGTGGAGCGGCTGCGAGCTTCGCGGATGG 626  
QY 813 CCGCGTATATGTCGAAACGTTGTGATTAAACCTCAGCATATTTAGTGAAGTGCAGATCTTGG 872  
Db 627 CCGAGTCTACGCGGAGCGTGCAGTAATCAATCCTCAGCATATTTAGGTCGAGATCTTGG 686  
QY 873 CGATCAGCTGAGAGTGTATACCTTTATGAAGCTGACTGCTCAGTGCAGCGTCTGCA 932  
Db 687 TGACTCCGTGCGCAACATCTCCACTGTACGAGGCTGATTTGCTGCTGCGAGCTCGCA 746  
QY 933 CCAAAAAGTTGTGAAAATTCGCCAGCAGCATTTGGATCCAGAACTGGCGTATCGCAT 992

Db 747 CCAAAAGGTCGTAGAGATCGCCACAGCACAGCACCTTTGGCGCCCGGACTACCGGAAAAGAT 806  
Qy 993 TTGTGGGATGACGTAAGATTTCTGCGCTCCATTGGTTACAGGGCGGGGAACGGTGA 1052  
Db 807 TTGTGACAGCGAGTCGGCTTTGCGCGCCACATCGGTTACCAAGGTGCGAGCACGGTGA 866  
Qy 1053 ATCTTTGGTCCGATGAAAAGGGCAACACAGCTTTCTCATCGAAATGAACCCACCGTATCCAGGT 1112  
Db 867 GTTCTAGTCGATGAGGAAGCAACACAGCTTTCTCATCGAGATGAACCTCGTATCCAGGT 926  
Qy 1113 TGAGCACACCGTGACTGAAGAAGTCAACGAGGTGACCTGTGAAGGGCGAGATGCGGTT 1172  
Db 927 CGAACACACCGGTGACCGAAGAGGTCAACCAAGTGCAGCTTGGTGAAGTCCCGACATCATAT 986  
Qy 1173 GCGTGTGGTCAACCTTTGAAGGAATGGGTCTGACCCAAAGATCAAGATCAAGACCCACGG 1232  
Db 987 CGCTTCGGTGGCACTCGAGATCTCGGTCTGGCCCAAGAGACATCCACCGAAGG 1046  
Qy 1233 TGAGGACATGAGTGGCGGATCACACGAGAGATCCAAACAGCGCTTCCGCCAGATAC 1292  
Db 1047 CGCTGCCCTGCGAGTGGCGTATTACTGAAGACCCCAACACCGTTCCTGCTGCAC 1106  
Qy 1293 CGGAATATCACCGCGTACCGCTCACAGGGGAGCTGGCGTTGGTCTTGACGGTGCAGC 1352  
Db 1107 CGGAAGATTAACCGCTTATCGTTCCCGAGGTGGCGAGGTGTTGGTCTCGAGCGTGTGC 1166  
Qy 1353 TCAGTCTGGTGGCGAAATCACCGCACATTTGACATCCATGCTGGTGAAGATGAACGTCGCG 1412  
Db 1167 CATGCTCGGTGGTGAATCACCCGAAATTTGACCTCCATGCTGGTCAAGATGAACGTCGCG 1226  
Qy 1413 TGGTTCGACTTTGAACCTGCTGCTGTCGTCAGACGCGCGTGGCTGGATTCACCGT 1472  
Db 1227 TGGTGGGATTTGCGCACCGGATGCTGCGCTCAGGTGCACTCGAGATTTGGTTGT 1286  
Qy 1473 GTCTGGTGTGCAACCAACATTTGGTTCTTGGTGGCTTGTGGGGAAGAGGACTTCAC 1532  
Db 1287 CTGAGGTGTGCTACCAACATCGGATTCCTGGTGGCTTGTACGCGAAGAGACTTCCA 1346  
Qy 1533 TTCCAGCGCATGCGCACCGGATTCATTGGCGATCACCGGACCTCTTCAGGCTCCACC 1592  
Db 1347 GCACAGCGCATGCGCACCTGGCTTATCGAGATCACCCATGGCTGCTGCGAGGACACC 1406  
Qy 1593 TGCTGATGATGACGAGGAGCGATCCTGGATTTACTTGGCAGATGTCAACGTTGAACAGCC 1652  
Db 1407 TGCCGACGATGAGCTGGCGGCAATTTGAACCTACCTAGCTGATGTGACTGTGAACAGCC 1466  
Qy 1653 TCATGGTGTGCTCAAGAGATGTTGCGAGCTCTCTATCGATAGCTGCTTAACATCAAGGA 1712  
Db 1467 TCATGGTCTACGCCCA--GCGGTGATCAACCCGCTAGAAAGCTCCAGCAGAGATCAA 1523  
Qy 1713 TCTGCCACTTGCCACCGGTTCCCGTGACCGCTGAAGCAGCTTTGGCCCGACGCGGTTTGC 1772  
Db 1524 AGGTGAGCTACCACTGGTTCCCGTGATCGCTGCTCCAGCTCGGCCAGAGGGATTTGC 1593  
Qy 1773 TCGTGAATCCGTGAGCAGGAGCGACTGGCGATTTACTGATACCACTTCCGCGATGACACA 1832  
Db 1584 GCGCGCACTGCGTAAACAGGATGCGCTCGCGTTACCGATACCACTTCCGTTGATGACA 1643  
Qy 1833 CGAGTCTTTGCTTGCACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGACAGGCGCT 1892  
Db 1644 TCAGTGGCTTCTCGCAACCCGCGTGGCTTCCAAACGCTGATCGAGCGCAGCACGCACT 1703  
Qy 1893 CGCAAGCTGACTCTCTGAGCTTTTGGTGGAGGCTGGGGCGGCGGACCTTAGATGT 1952  
Db 1704 CGCAAGCTCACCCACAGCTGCTCTCTGTAGAAGCATGGGTGGTGTCTACCTATGAGCT 1763  
Qy 1953 GCGATGCGGTTTCTCTTTGAGATCCGTTGGGACAGGCTTCGACGAGCTGCGCGAGGCGAT 2012  
Db 1764 TGCTATGCGCTTCTTACATAGAGGCCCTTGGAGAGGCTCTCGACACCTTGGGTGAGGCCAT 1823  
Qy 2013 GCGAATGAAACATTCAGATGCTGTTTCGCGCGCGCAACACCGTGGGATACACCCCGTA 2072  
Db 1824 GCGCAAGCTCAACATTCAGATGCTGTCGAGGCGCGCAACACAGTCCGCTACACACCGTA 1883

Qy 2073 CCGAGACTCCGTCTCGCGCGCTTTGTTAAGGAAGTCCAGCTCGCGGTGACATCTT 2132  
Db 1884 CCGAGATTCGTATGTCAGGCTTTTGTGCAAGCGCCACGCTCTGGTGTGATATCTT 1943  
Qy 2133 CCGCATTTTCGACGGCTTTAAGCAGCTCTCCAGATGCGTCCAGCAATCGACGAGTCT 2192  
Db 1944 CCGCATTTTCGACGGCTTTAAGCAGCTCTCCAGATGCGTCCAGCAATCGACGAGTCT 2003  
Qy 2193 GGNACCAACCGGGTAGCGAGGTGGCTATGCTTATTTCTGGTGAATCTCTGATCC 2252  
Db 2004 CGAAACCAACACACGANTTCCGGAATTCGAATGGCTTACTCCGGTGAATCGCAACCC 2063  
Qy 2253 AATGAAAAGCTCTACACCCCTGGATTTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTC 2312  
Db 2064 AAGTGAGAAGCTCTACACACTGGATTTACTACCTTAAGCTTTCGAGGAATCGTTAAGTC 2123  
Qy 2313 TGGCGCTCAATCTTGGCCATTAAAGATATGCTGGTCTGCTTCGCCACGCTCGGTAAC 2372  
Db 2124 CCGTGCACATGTTCTTAGCCATTAAAGACATGCTGACATGAAAGCCAGCCGCGACAC 2183  
Qy 2373 CAAGCTGGTCAACCGCACTCGCGCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGA 2432  
Db 2184 CAAGCTGGTCAACCGCACTAGAGCGCAACTTCGACCTGCGGTTTCATGTTTCAACCCACGA 2243  
Qy 2433 CACTCGGGTGGCAGCTGGCAACCTACTTTGCTGAGCTCAAGCTGGTGGAGATGCTGT 2492  
Db 2244 CACTGCTGGTGGCCAGCTGGCTACGTTATTTGGCTGCGAGCGCTGGTGGAGATGCGCT 2303  
Qy 2493 TGACGCTGTCTTCGCAACCACTGTCTGGCACCACTTCCAGCCATCCCTGCTCTGCCATTTG 2552  
Db 2304 CGATGCTGTCTTCGCGCGCTGCTGGGCAACAGTTCGAGCCATCGTGTCTGCAATTTG 2363  
Qy 2553 TGCTGATTCGCGCACACCCGTCGCGATACCGGTTTGAGCTGCGAGGCTGTTTCTGACCT 2612  
Db 2364 GGCAGCATTCGCTTAAACCTACCGGATACCGGCTCTCTCTGCGAGCTCGGATCGAT 2423  
Qy 2613 CGAGCGCTACTGGGAAGCAGTGGCGGACTGTACCTGCTCCAGCCATCCCTGCTGGAACCCACG 2672  
Db 2424 GGAACGCTACTGGGAAGCAGTGGGCAACTTACGCACTTCGATCCGGAACCCACG 2483  
Qy 2673 CCAACCGGTGCGCTTACCGGCAACAAATCCAGCGGACAGTGTTCACCTGCGTGC 2732  
Db 2484 ACCAATGCGCGGCTTACCGACCAAGAAATCCAGCGGCGGAGCTGTCCAATTCGCGCG 2543  
Qy 2733 ACAGGCAACGCACTGGGCTTTCGAGATCGTTTCGAATCATCGAGAGCAATACGACG 2792  
Db 2544 ACAGGCAACGCACTGGGCTTTCGAGATCGTTTCGAGCTCATCGAGGATCTTACGCG 2603  
Qy 2793 CGTTAATGAGATGCTGGGACGCGCCAAACCAAGGTCAACCCATCCTCCAAGTGTTCGCGA 2852  
Db 2604 CGTTAATGAGATGCTGGGACGCGCCCAACCAAGGTCAACCCGCTCTCAAGGTTCGCGA 2663  
Qy 2853 CTTGCACTCCACTGTTGGTGGGCTGTGATCCAGCAGACTTTGCTGCGGATCCACA 2912  
Db 2664 TCTTGCACCTTACTAGTGGGCGCAGAGTTAACCCAGCGGACTTTTCAGAGATCCCA 2723  
Qy 2913 AAAGTACGACATCCCGACATCTGTCATCGGTTTCTTCGCGCGGAGCTTTGGTAACCCCTC 2972  
Db 2724 AAAGTACGATTTCTTGAATCTCGCTCATTTGCTTCTTCGTTGGTGGTGGCTCGGACCCCTC 2783  
Qy 2973 AGTGGCTGGCAGAGCACTGCGCACCGCGCAGCTGGAAGGCGCTCCGAGGCGAGGC 3032  
Db 2784 AGCGGCTGGCCAGAGAAATGCGCTGTCTAAGCATTTGGCAGGACGTAAGGAATCAAGGA 2843  
Qy 3033 ACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGAGCTGATGATTTCCAAGGA 3092  
Db 2844 CACATTTGGCCCCGCTGCCAGCAGAAAGATGAAGCTCTGCTCGCAGATCGTGGGAGCC 2898  
Qy 3093 AGTTCGCAATAGGCTCAACCGCTGCTGTTCCGGAAGCCAAACGGAAGGTTCTTCAGCA 3152  
Db 2899 -GTTCCGCAACAGCTGATGCTGCTCTTCTCCGGAAGCCCTGGCGGCAATTCGAGAGCA 2957

```
QY 3153 CGGTGCGCCGCTTCGGCAACACACTCTCGCTGGATGATCGTGAATTCCTACGGCCTGGT 3212
Db 2958 CCGCGCTCAGTTTGGCGATACCAACCAACTCGGTGACGAGAAATTCCTCTATGGCCTCAA 3017
QY 3213 CGAAGCCGCGAGACTTTGATCGGCTGCGCAGATGGCGACCCCACTGCTTGTGCGCT 3272
Db 3018 AGAAGCAAGAAACGGTCAATTCGTACCGCAGACAGTAGTGCCCAATGCTGTGCGTCT 3077
QY 3273 GATGCGATCTCTGAGCCAGACGATAAGGGTATGGGCAATGTTGTGGCAACGTCACGG 3332
Db 3078 CGATGCGCTCGCGAGCCCGACGAAAGGCGATGGCAACGTTGTGTCAATGTCAACGG 3137
QY 3333 CGAGTCGCCCACTCGTGTGGTGACCGCTCCGTTGAGTGTGTCACGCAACCGCAGA 3392
Db 3138 CCAGATTCGCCCATCTCTGTCGGTATCGCAGCGTAGTCAAGTACGTCGCTCGGTGGA 3197
QY 3393 AAAGGCAGATTCCTCCAAACAGGCGCATGTTGTCGACCAATCGCTGTTGTCAACGT 3452
Db 3198 AAAGGCTGATCTCTTCAACCGCGACACGTAGCAGCACCATTCGCGAGCGTTGTCAACGT 3257
QY 3453 GACTGTTGCTGAGGTGATGAGTCAAGGCTCAAGCTGAGTGCAGTCGCAATCATCGAGCTAT 3512
Db 3258 GACAATCGAAGAGCGCTACCGTCAAGCGCGGATCCAGTTCGTTGTTGAGCGAAT 3317
QY 3513 GAAGATGGAAGCAACAACTCACTCTTCTGTGTGACGGCAAAATCGATCGCGTTGTGCTTCC 3572
Db 3318 GAAGATGGAAGCAACGATCTCTGCCACCACAGAGGGAACCGTCGACCGTATCGTCTTGAC 3377
QY 3573 TGTGTCACGAAGTGGAGGTGGCACTGTGATCGTCT 3611
Db 3378 GCAGGCCACCAGGTGGAAGGTGGAGACTTGCTCTCGT 3416

RESULT 11
AAF71419
ID AAF71419 standard; DNA; 1719 BP.
AC AAF71419;
XX
XX
XX 30-APR-2001 (first entry)
XX
XX Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:119.
XX
XX Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
XX
OS Corynebacterium glutamicum.
XX
XX WO200100844-A2.
XX
XX 04-JAN-2001.
XX
XX 23-JUN-2000; 2000WO-IB000943.
XX
XX 25-JUN-1999; 99US-0141031P.
XX 08-JUL-1999; 99DE-01031412.
XX 08-JUL-1999; 99DE-01031413.
XX 08-JUL-1999; 99DE-01031419.
XX 08-JUL-1999; 99DE-01031420.
XX 08-JUL-1999; 99DE-01031420.
XX 08-JUL-1999; 99DE-01031424.
XX 08-JUL-1999; 99DE-01031428.
XX 08-JUL-1999; 99DE-01031431.
XX 08-JUL-1999; 99DE-01031431.
XX 08-JUL-1999; 99DE-01031433.
XX 08-JUL-1999; 99DE-01031434.
XX 08-JUL-1999; 99DE-01031510.
XX 08-JUL-1999; 99DE-01031562.
XX 08-JUL-1999; 99DE-01031634.
XX 09-JUL-1999; 99DE-01032180.
```

```
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99US-0143208P.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033005.
PR 27-AUG-1999; 99DE-01040765.
PR 31-AUG-1999; 99US-0151572P.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042123.
PR 03-SEP-1999; 99DE-01042125.
XX (BADI ) BASF AG.
XX
XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
XX P-PSDB; AAB79302.
XX
XX WPI; 2001-061975/07.
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
XX metabolism and oxidative phosphorylation protein for production or
XX modulation of production of fine chemicals e.g. amino acids,
XX carbohydrates or enzymes.
XX
XX Claim 3; Page 316-319; 1246pp; English.
XX
XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
XX metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
XX to AAB 79633 which are involved in carbon metabolism and energy
XX production. The C. glutamicum SMP gene can be used in vectors (II) for
XX expression in host cells and production or modulation of production of
XX fine chemicals, such as, an organic acid, a proteinogenic or
XX nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
XX nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
XX a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
XX polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
XX encoded by them are used for diagnosing the presence or activity of
XX Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
XX containing them are used to map genomes of organisms related to C.
XX glutamicum, identify and localise C. glutamicum sequences of interest, in
XX evolutionary studies, in determining SMP protein regions required for
XX function, in modulating SMP protein activity, in modulating the
XX metabolism of sugars, and in modulating high-energy molecule production
XX in a cell (i.e. ATP, NADPH)
XX
XX Sequence 1719 BP; 405 A; 457 C; 468 G; 389 T; 0 U; 0 Other;
XX
XX Query Match 47.2%; Score 1709; DB 4; Length 1719;
XX Best Local Similarity 99.9%; Pred.No. 0;
XX Matches 1719; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 48 GAAAGTCGTGCAGGTCAGGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAAAACCGAT 107
Db 1 GAAGTCGTGCAGGTCAGGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAAAACCGAT 60
QY 108 GTTTGATTGGGGGAATCGGGGTTACGATAGTAGGAGCGAGTGTATCACCCTTGG 167
Db 61 GTTTGATTGGGGGAATCGGGGTTACGATAGTAGGAGCGAGTGTATCACCCTTGG 120
QY 168 CGGTCTCTTTGTTGAAAGGAATAATTAATCTAGTGTGCACTCACACATCTTCAACGCTTCC 227
Db 121 CGGTCTCTTTGTTGAAAGGAATAATTAATCTAGTGTGCACTCACACATCTTCAACGCTTCC 180
QY 228 AGCATTTCAAAAGATCTTTGGTAGCAAAACCGCGGGAATTCGGTTCGGTTCGGTGC 287
Db 181 AGCATTTCAAAAGATCTTTGGTAGCAAAACCGCGGGAATTCGGTTCGGTTCGGTGC 240
QY 288 AGCACTCGAAACCGGTGTCAGCCACCGTAGTATTTACCCCGTGAAGATCGGGATCAT 347
```

Db 241 AGCACTCGAAACCGGTGACGACACGGTAGCTATTTACCCCGGTGAAGATCGGGGATCAAT 300  
QY 348 CCACCGCTCTTTTGTCTTGAAGCTGTCCGCAATGGTACCGAAGGCTCACCAAGTCAAGGC 407  
Db 301 CCACCGCTCTTTTGTCTTGAAGCTGTCCGCAATGGTACCGAAGGCTCACCAAGTCAAGGC 360  
QY 408 GTACTGTGACATCGATGAATTTATCGGTGACGCTTAAAGATTAAGCAGATGCCATTTA 467  
Db 361 GTACTGTGACATCGATGAATTTATCGGTGACGCTTAAAGATTAAGCAGATGCCATTTA 420  
QY 468 CCGGGATACGGCTCTCTGTGAATAAGCCAGCTTGGCCGAGTGTGGGAAAACGG 527  
Db 421 CCGGGATACGGCTCTCTGTGAATAAGCCAGCTTGGCCGAGTGTGGGAAAACGG 480  
QY 528 CAATTAATTTTATGGCCAAACCCAGAGGTTCTTGATCTCACCGTGAATAGTCTCGGC 587  
Db 481 CAATTAATTTTATGGCCAAACCCAGAGGTTCTTGATCTCACCGTGAATAGTCTCGGC 540  
QY 588 GGTAAACCGCGGAGAAAGGCTGTGCTGCAAGTTTGGCGGAATCCACCCGAGCAAAA 647  
Db 541 GGTAAACCGCGGAGAAAGGCTGTGCTGCAAGTTTGGCGGAATCCACCCGAGCAAAA 600  
QY 648 CATCGATGAGATCGTTTAAAGCGCTGAAGCCAGACTTACCCATCTTTTGTGAAGGAGT 707  
Db 601 CATCGATGAGATCGTTTAAAGCGCTGAAGCCAGACTTACCCATCTTTTGTGAAGGAGT 660  
QY 708 TCCCGGTGTGGCGGACGGGTATGCTGTTTGTGCTTCCATCTGATGAGTTCGCAAT 767  
Db 661 TCCCGGTGTGGCGGACGGGTATGCTGTTTGTGCTTCCATCTGATGAGTTCGCAAT 720  
QY 768 AGCAACAGAAAGCATCTCGTGAAGCTGAAGCGCTTTCGGCGATGGCGGATATGTGCA 827  
Db 721 AGCAACAGAAAGCATCTCGTGAAGCTGAAGCGCTTTCGGCGATGGCGGATATGTGCA 780  
QY 828 ACCTGCTGTGATTAACCCCTCAGATATTAAGTGCAGATCCCTTGGCGATCACTTGGAGA 887  
Db 781 ACCTGCTGTGATTAACCCCTCAGATATTAAGTGCAGATCCCTTGGCGATCACTTGGAGA 840  
QY 888 AGTTGTACACCTTTATGAAGCTGACTGCTCACTGACGCTGCTGCAAAAAGTTGTGCA 947  
Db 841 AGTTGTACACCTTTATGAAGCTGACTGCTCACTGACGCTGCTGCAAAAAGTTGTGCA 900  
QY 948 AATTGGCCAGCAGCATTTGATGATCCAGAACTGGCTGATGATGCTGCTGCTGCTGCTGCT 1007  
Db 901 AATTGGCCAGCAGCATTTGATGATCCAGAACTGGCTGATGATGCTGCTGCTGCTGCTGCT 960  
QY 1008 AAAGTTCTGCGCTCCATTTGTTTACAGGCGCGGGAACCGTGGAAATCTTGTGCTGATGA 1067  
Db 961 AAAGTTCTGCGCTCCATTTGTTTACAGGCGCGGGAACCGTGGAAATCTTGTGCTGATGA 1020  
QY 1068 AAAGGGAACCAACGCTTTCATCGAATGAACCCACGATCCAGGTTGAGCAACCGTGAC 1127  
Db 1021 AAAGGGAACCAACGCTTTCATCGAATGAACCCACGATCCAGGTTGAGCAACCGTGAC 1080  
QY 1128 TGAAGAAGTCAACGAGTGACCTGTGAGGCGCAGATGCGCTTGGCTGCTGCTGCTGCTGCT 1187  
Db 1081 TGAAGAAGTCAACGAGTGACCTGTGAGGCGCAGATGCGCTTGGCTGCTGCTGCTGCTGCT 1140  
QY 1188 CTTGAAGAAATGGGTCTGACCCCAAGATGAATGAAGATCAAGACCCAGTGCAGAGT 1247  
Db 1141 CTTGAAGAAATGGGTCTGACCCCAAGATGAATGAAGATCAAGACCCAGTGCAGAGT 1200  
QY 1248 CCGCATCACCAGGAGATCCAAACAGCGCTTCCGCCAGATACCGGAACCTATCACCGC 1307  
Db 1201 CCGCATCACCAGGAGATCCAAACAGCGCTTCCGCCAGATACCGGAACCTATCACCGC 1260  
QY 1308 GTACCGCTCACAGGCGGAGCTGGCGTTCTGTGACGCTGACGCTCAGCTCGGTGGCA 1367  
Db 1261 GTACCGCTCACAGGCGGAGCTGGCGTTCTGTGACGCTGACGCTCAGCTCGGTGGCA 1320  
QY 1368 AATCACCGCACATTTGACTCCATGCTGTTGAATGAATGACCTGCTGCTGCTGCTGCTGCTGCT 1427

Db 1321 AATCACCGCACATTTGACTCCATGCTGCTGTAAGATGACCTGCCGTGGTTCGACTTTGA 1380  
QY 1428 AATGCTGTTGCTGCTGTCACAGCGCGCTTGGCTGAGTTTCAAGTGTCTGGTGTGCAAC 1487  
Db 1381 AATGCTGTTGCTGCTGTCACAGCGCGCTTGGCTGAGTTTCAAGTGTCTGGTGTGCAAC 1440  
QY 1488 CAACATTTGGTTTCTTGGCTGCTGCTGCGGAAGAGGACTTCACTTCCAAGCGCATCGC 1547  
Db 1441 CAACATTTGGTTTCTTGGCTGCTGCTGCGGAAGAGGACTTCACTTCCAAGCGCATCGC 1500  
QY 1548 CACCGATTCATTCGCGATCACCCGACCTCTTCAAGGCTCAACCTGCTGATGATGAGCA 1607  
Db 1501 CACCGATTCATTCGCGATCACCCGACCTCTTCAAGGCTCAACCTGCTGATGATGAGCA 1560  
QY 1608 GCGACGATCTCTGATTTACTTGGCAGATGTCAAGTGAACAGCCCTCATGGTGTGCGTCC 1667  
Db 1561 GCGACGATCTCTGATTTACTTGGCAGATGTCAAGTGAACAGCCCTCATGGTGTGCGTCC 1620  
QY 1668 AAAGGATTTGAGCTCTTATGATTAAGTGCCTTAACATCAAGGATCTGCACTGCCACG 1727  
Db 1621 AAAGGATTTGAGCTCTTATGATTAAGTGCCTTAACATCAAGGATCTGCACTGCCACG 1680  
QY 1728 CGGTTCCCGTGACGCTGAGCAGCTTGGCCCGCGCG 1767  
Db 1681 CGGTTCCCGTGACGCTGAGCAGCTTGGCCCGCGCG 1719

RESULT 12  
AAF71420  
ID AAF71420 standard; DNA; 1406 BP.  
XX  
AC AAF71420;  
XX  
DT 30-APR-2001 (first entry)  
XX  
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:121.  
XX  
KW Corynebacterium glutamicum; carbon metabolism and energy production;  
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;  
KW fine chemical production; organic acid; proteinogenic amino acid;  
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;  
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;  
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;  
XX diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.  
OS Corynebacterium glutamicum.  
XX  
PN WO200100844-A2.  
XX  
PD 04-JAN-2001.  
XX  
PF 23-JUN-2000; 2000WO-IB000943.  
XX  
PR 25-JUN-1999; 99US-0141031P.  
PR 08-JUL-1999; 99DE-01031412.  
PR 08-JUL-1999; 99DE-01031413.  
PR 08-JUL-1999; 99DE-01031419.  
PR 08-JUL-1999; 99DE-01031420.  
PR 08-JUL-1999; 99DE-01031424.  
PR 08-JUL-1999; 99DE-01031428.  
PR 08-JUL-1999; 99DE-01031431.  
PR 08-JUL-1999; 99DE-01031433.  
PR 08-JUL-1999; 99DE-01031434.  
PR 08-JUL-1999; 99DE-01031510.  
PR 08-JUL-1999; 99DE-01031512.  
PR 08-JUL-1999; 99DE-01031562.  
PR 09-JUL-1999; 99DE-01032180.  
PR 09-JUL-1999; 99DE-01032227.  
PR 09-JUL-1999; 99US-0143208P.  
PR 14-JUL-1999; 99DE-01032924.  
PR 14-JUL-1999; 99DE-01032973.  
PR 14-JUL-1999; 99DE-01033005.



PR 27-AUG-1999; 99DE-01040765.  
PR 31-AUG-1999; 99US-0151572P.  
PR 03-SEP-1999; 99DE-01042076.  
PR 03-SEP-1999; 99DE-01042079.  
PR 03-SEP-1999; 99DE-01042086.  
PR 03-SEP-1999; 99DE-01042087.  
PR 03-SEP-1999; 99DE-01042088.  
PR 03-SEP-1999; 99DE-01042095.  
PR 03-SEP-1999; 99DE-01042123.  
PR 03-SEP-1999; 99DE-01042125.  
XX (BADI ) BASF AG.  
XX  
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Habershauer G;  
XX  
XX WPI; 2001-061975/07.  
XX P-ESDB; AAB79303.  
XX  
XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar  
XX metabolism and oxidative phosphorylation protein for production or  
XX modulation of production of fine chemicals e.g. amino acids,  
XX carbohydrates or enzymes.  
XX  
XX Claim 3; Page 321-323; 1246pp; English.  
XX  
XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar  
XX metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243  
XX to AAB 79633 which are involved in carbon metabolism and energy  
XX production. The C. glutamicum SMP gene can be used in vectors (II) for  
XX expression in host cells and production or modulation of production of  
XX fine chemicals, such as, an organic acid, a proteinogenic or  
XX nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a  
XX nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,  
XX a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a  
XX polyketide, or an enzyme. The presence of (I) or SMP proteins (III)  
XX encoded by them are used for diagnosing the presence or activity of  
XX Corynebacterium diphtheriae in a subject. (II), (II), (III) or host cells  
XX containing them are used to map genomes of organisms related to C.  
XX glutamicum, identify and localise C. glutamicum sequences of interest, in  
XX evolutionary studies, in determining SMP protein regions required for  
XX function, in modulating SMP protein activity, in modulating the  
XX metabolism of sugars, and in modulating high-energy molecule production  
XX in a cell (i.e. ATP, NADPH)  
XX  
XX Sequence 1406 BP; 325 A; 381 C; 385 G; 315 T; 0 U; 0 Other;  
XX  
XX Query Match 38.5%; Score 1395; DB 4; Length 1406;  
XX Best Local Similarity 99.9%; Pred. No. 0;  
XX Matches 1406; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
XX  
QY 361 GCTTCTGAAGCTGTCGCGATTTGGTACCGAGGCTCACCGAGGTACCTGGACATC 420  
Db 1 GCTTCTGAAGCTGTCGCGATTTGGTACCGAGGCTCACCGAGGTACCTGGACATC 60  
QY 421 GATGAATTTATCGGTGCGATTAAGGTTAAAGTAAAGATGATGCAATTTACCGGGATCGGC 480  
Db 61 GATGAATTTATCGGTGCGATTAAGGTTAAAGTAAAGATGATGCAATTTACCGGGATCGGC 120  
QY 481 TTCCTGTGTAATGTCGCGATTTGGTACCGAGGCTCACCGAGGTACCTGGACATC 540  
Db 121 TTCCTGTGTAATGTCGCGATTTGGTACCGAGGCTCACCGAGGTACCTGGACATC 180  
QY 541 GGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATTAAGTCTCGCGGGTAAACCGCGCG 600  
Db 181 GGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATTAAGTCTCGCGGGTAAACCGCGCG 240  
QY 601 AAGAGGCTGTCGCGATTTGGGGAATTCACCGGACGCAACCAATCGATGATC 660  
Db 241 AAGAGGCTGTCGCGATTTGGGGAATTCACCGGACGCAACCAATCGATGATC 300  
QY 661 GTTAAAGCGGTGAAGGCCAGCATTTACCCATCTTTGTGAAGGCGATTTGCCGGTGGTGGC 720  
Db 301 GTTAAAGCGGTGAAGGCCAGCATTTACCCATCTTTGTGAAGGCGATTTGCCGGTGGTGGC 360

QY 721 GGACGCGGTATGCGTTTGTGTTCTTACCTGATGAGCTTCGCAAAATTAGCAACAGAGCA 780  
Db 361 GGACGCGGTATGCGTTTGTGTTCTTACCTGATGAGCTTCGCAAAATTAGCAACAGAGCA 420  
QY 781 TCTGTTGAAGCTGAAGCGGCTTTCCGGCGATGGCGCGTATATGTCGAACGTCGTGTGATT 840  
Db 421 TCTGTTGAAGCTGAAGCGGCTTTCCGGCGATGGCGCGTATATGTCGAACGTCGTGTGATT 480  
QY 841 AACCTCTAGCATATTAAGTGCAGATCCTTGGCGATCACAATGAGAGAGTTGTACACTT 900  
Db 481 AACCTCTAGCATATTAAGTGCAGATCCTTGGCGATCACAATGAGAGAGTTGTACACTT 540  
QY 901 TATGAACGTGACTGCTCACTGACGCTGCTCACCAAAAAGTTGCGAAAATTGCGCCAGCA 960  
Db 541 TATGAACGTGACTGCTCACTGACGCTGCTCACCAAAAAGTTGCGAAAATTGCGCCAGCA 600  
QY 961 CAGCATTTGGATTCAGAACTGCGGTGATGCGANTTTGTCGGATGCGATGAAGTTGTGCGGC 1020  
Db 601 CAGCATTTGGATTCAGAACTGCGGTGATGCGANTTTGTCGGATGCGATGAAGTTGTGCGGC 660  
QY 1021 TCCATTGTTTACCAGGCGCGGAAACGTCGAAATCTTGGTTCGATGAAGAGGCAACAC 1080  
Db 661 TCCATTGTTTACCAGGCGCGGAAACGTCGAAATCTTGGTTCGATGAAGAGGCAACAC 720  
QY 1081 GTCTTCATCGAAATGAACCCACGATATCAGGTTGAGCACACCGTGTGATGAAGAGTCAAC 1140  
Db 721 GTCTTCATCGAAATGAACCCACGATATCAGGTTGAGCACACCGTGTGATGAAGAGTCAAC 780  
QY 1141 GAGGTGACCTGTTGAGGCGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
Db 781 GAGGTGACCTGTTGAGGCGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 1201 GGTTCGACCCAGATTAAGATCAAGACCCACCGTGCAGCATCTGCAGTGCCTGATCACAACG 1260  
Db 841 GGTTCGACCCAGATTAAGATCAAGACCCACCGTGCAGCATCTGCAGTGCCTGATCACAACG 900  
QY 1261 GAAGATCAACACAGGCTTCCGCGCGAGATACCGAACTATCACCGGTACCGTCAACA 1320  
Db 901 GAAGATCAACACAGGCTTCCGCGCGAGATACCGAACTATCACCGGTACCGTCAACA 960  
QY 1321 GGCGAGCTGCGGCTTGTGCTTTCGAGGTCAGCTCAGCTCGGTGCGGAAATCACCGCACAC 1380  
Db 961 GGCGAGCTGCGGCTTGTGCTTTCGAGGTCAGCTCAGCTCGGTGCGGAAATCACCGCACAC 1020  
QY 1381 TTTGATCTCATGCTGTTGAAATGACCTGCGGTGCTCCGACTTTGAACTGCTGTTGCT 1440  
Db 1021 TTTGATCTCATGCTGTTGAAATGACCTGCGGTGCTCCGACTTTGAACTGCTGTTGCT 1080  
QY 1441 CGTGCACAGCGCGCTTGGCTGAGTTTCAACCGTGTCTGCTGTTGCAACCAATTTGTTTC 1500  
Db 1081 CGTGCACAGCGCGCTTGGCTGAGTTTCAACCGTGTCTGCTGTTGCAACCAATTTGTTTC 1140  
QY 1501 TTGCGTGGTGTGCTGCGGGAAGAGACTTCACTTCCAGCGCATCGCCACCGGATTCATT 1560  
Db 1141 TTGCGTGGTGTGCTGCGGGAAGAGACTTCACTTCCAGCGCATCGCCACCGGATTCATT 1200  
QY 1561 GGCGATCACCCGCACTCTCTTCCAGGCTCCACTGCTGATGATGAGCAGGAGCATCTCTG 1620  
Db 1201 GGCGATCACCCGCACTCTCTTCCAGGCTCCACTGCTGATGATGAGCAGGAGCATCTCTG 1260  
QY 1621 GATTAATTTGGCAGATGTCACCGTGAACAGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
Db 1261 GATTAATTTGGCAGATGTCACCGTGAACAGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1681 GCTCTATTCGATAAGCTGCTTAAATCAAGGATCTGCCACTTGCACCGCGGTTCCCGGTGAC 1740  
Db 1321 GCTCTATTCGATAAGCTGCTTAAATCAAGGATCTGCCACTTGCACCGCGGTTCCCGGTGAC 1380  
QY 1741 CGCTGTAAGCAGCTTGGCCCGACCGCGC 1767  
Db 1381 CG-CTGAAGCAGCTTGGCCCGACCGCGC 1406

RESULT 13

ACA37896

ID ACA37896 standard; DNA; 3381 BP.

XX AC

XX AC

XX AC

XX 19-JUN-2003 (first entry)

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #19553.

XX Antisense; db; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

XX Mycobacterium avium.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABU34026.

XX New antisense nucleic acids, useful for identifying proteins or screening

XX for homologous nucleic acids required for cellular proliferation to

XX isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 25766; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

XX the 6213 antisense sequences given in the specification where expression

XX of the nucleic acid inhibits proliferation of a cell. Also included are:

XX (1) a vector comprising a promoter operably linked to the nucleic acid

XX encoding a polypeptide whose expression is inhibited by the antisense

XX nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX polypeptide or its fragment whose expression is inhibited by the

XX antisense nucleic acid; (4) an antibody capable of specifically binding

XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX proliferation or the activity of a gene in an operon required for

XX proliferation; (7) identifying a compound that influences the activity of

XX the gene product or that has an activity against a biological pathway

XX required for proliferation, or that inhibits cellular proliferation; (8)

XX identifying a gene required for cellular proliferation or the biological

XX pathway in which a proliferation-required gene or its gene product lies

XX or a gene on which the test compound that inhibits proliferation of an

XX organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX product is overexpressed or underexpressed; (12) determining the extent

XX to which each of the strains is present in a culture or collection of

XX strains; or (13) identifying the target of a compound that inhibits the

XX proliferation of an organism. The antisense nucleic acids are useful for

XX identifying proteins or screening for homologous nucleic acids required

XX for cellular proliferation to isolate candidate molecules for rational

XX drug discovery programs, or for screening homologous nucleic acids

XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

XX *k. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

XX not form part of the printed specification, but was obtained in

XX electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences

XX	Sequence	3381 BP; 542 A; 1154 C; 1196 G; 489 T; 0 U; 0 Other;
SQ	Query Match	37.6%; Score 1361.6; DB 7; Length 3381;
	Best Local Similarity	63.2%; Pred. No. 0;
	Matches 2131; Conservative	0; Mismatches 1234; Indels 9; Gaps 2;
QY	238	AAGATCTTGGTAGCAAAACCGGGCGAATCGCGTTCGGTTCCTTCCGTGAGCACTCGAA 297
Db	10	AAAGTCTCTGGTTCGCCAACCGCGGGGAGATCGGATCCGCGGTTCCCGCGGCTACGAG 69
QY	298	ACCGGTGAGCAGCAGGTAGTATTATACCCCGTGAAGATCGGGATCATTTCCACCGTCT 357
Db	70	CTGGAATGGCCACCGTGGGTGTATCCCTACGAGACCGCAATTCGGTGACCGGTG 129
QY	358	TTTGCTTCTGAAGCTGTCCGATTTGGTACCGAAGGCTCACGAGTCAAGGGTCACTCGAC 417
Db	130	AAGCCGACGAGTCTTACAGATCGGTGAGGAGGCCACCCAGTCCGCGCTTCTCTTCC 189
QY	418	ATCGATGAATTTATCGGTGCAGCTAAAGATTTAAAGCAGATGCCATTTACCCGGGATAC 477
Db	190	GTCCGACGAGATCGTTCGGCACCGCGCTGCGCGCGGATGCGATCTACCCCGGCTAC 249
QY	478	GGCTTCTGTCTGAAATGCCAGCTTCCCGCGAGTGTGCGGAAAGCGGATCTTTT 537
Db	250	GGCTTCTGTCTGCGGAAACCCCGATCTGCGCGCGGTGCGCGCGCGGCGCATCACGTT 309
QY	538	ATTGGCCCAACCCAGAGGTTCTTGATCTCACCGTGATAAGTCTCGCGCGGTAAACGCC 597
Db	310	GTCCGCCCCAGCGCGAGTGTCTGAGTCTACCGCGCAAGTCCGCGGCCCATCGCGCG 369
QY	598	GGGAGAGAGGTGTCTCCAGTTTGGCGGAATCCACCCGAGCAAAACATCGATGAG 657
Db	370	GGCGGGCGCGGGTCTTCCGGTGTGCGCGCTCTCGCGCGCTCGACCTCGGTGCGAGAG 429
QY	658	ATCGTTAAAGCGCTGAAGCGCAGACTTACCCCATCTTTGTGAAGCGAGTTGCCGGTGT 717
Db	430	TTGCTGAGCGCGCGGAGAGTACGTTCCGCTGTTCTGTCAGCGCGTTCGCGCGGCG 489
QY	718	GGCGGACGCGGTATCGTTTGTCTTCACTGATGAGCTTCGCAATTTAGCAACAGAA 777
Db	490	GGCGGGCGGGGATCGCGCGGTCAACCGATCCGGGGGCGCTGCGCGGAGCGATCGAGGCG 549
QY	778	GCATCTCGTGAAGCTGAAGCGCTTTCGGCGATGCGGGTATATGTGAAACGTCTGTG 837
Db	550	GCAGCGGTGAGCGCGAGTCCCGGTTCGGGAGCGCTCGGTGTTCTCGAGCAGCGGTG 609
QY	838	ATTAAACCTTCAGCATATTGAAGTGCAGATCTTGGCGATTCACACTGGAGAACTTGTAAC 897
Db	610	ATCAATCCCGGCGACATCGAGTCCAGATCTCTGGCGCGACACGCGCGCAACCTGATGCAC 669
QY	898	CTTTATGAAGTGAAGTGTCTACCTGAGCGTCTGTCACCAAAAAGTTGTGAAATTCGCCA 957
Db	670	CTCTACGAGCGGACTGCGAGCGTGCAGCGCGCGGACCAAGAGGTATCGAGTTCGCCCCG 729
QY	958	GCACAGCATTTGATCCAGAACTCGGTGATCGCATTTGTGCGGATGCACTAAAGTTCTGC 1017
Db	730	GGCGCCAACTTGATCCGGGCTGCGCGAACGATCTGCGCGACGCGGTGGGTTTCGCA 789
QY	1018	CGTTCATTTGTTTACGAGGCGCGGAAACGTTGGAATTTCTTGGTGCATGAAAGGCGAAC 1077
Db	790	CGCAGCATCGGTATCACCTGCGCGGCGACTGTGGAGTTCTCTGCTCGAACCGCGGCAAC 849
QY	1078	CAGTCTTTCATCGAAATGAACCCATCTATCCAGTTTGAGCACACCGTGTACTGAAGATTC 1137
Db	850	CAGTGTTCATCGAGTGAACCCCGCATCCAGTGTGAGCACACGTCGCGGAGATC 909
QY	1138	ACCGAGTGAAGTGTGAGGCGGAGATCGGTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
Db	910	ACCGACGTCGACCTGTGTTCGCGCGAGTGTGCGCATCGCTCCGCGCGAGAGTGTGAGAG 969
QY	1198	TTGGTCTGACCAAGATAGTCAAGACCCAGGTGCGAGTGTGCGGATGCGGATGCGGATCACC 1257

970 ATTGGCTGAGCCAGGATTCGGTCACTCTCA CGGCGCGCCCTGCA GTGCGGATCACC 1029 Db  
1258 ACGAAGATCCAAACAAAGCTTCGCCCCAGATACCGAATCATCACCGGTACCGCTCA 1317 Qy  
1030 ACCGAGGACCCGCCAAACGGTTTCGCCCCGGAACACCGGCCGATCACCGCTACCGCAC 1089 Db  
1318 CCAGGCGGAGCTGCGCTTCGTCTTGACGGTGAGCTTCAGCTCGGTGGGGAATACACGCA 1377 Qy  
1090 CCGGGTGGCGCGGATCCGGCTGAGACGGCGGACCAACCGCTGGCGCGGAGATCAGCGG 1149 Db  
1378 CACTTTGACTCCATGCTGGTGAATAGCTCGCGTGGTTCGACTTTGAACTGCTGTT 1437 Qy  
1150 CACTTCGATTCGATGCTGATCAAGCTGACTGCGGGGACGCGACTTCCCAACGGCGGTG 1209 Db  
1438 GCTCGTGCA CAGCGCGCTTGGCTGAGTTCAACCGTGTCTGCTGTTGTCACCAACCAATTGGT 1497 Qy  
1210 CGCGCGCGCGCGCGGTCGCGGAATTCGATCCGCGGGTCTCGACGAATATCCG 1269 Db  
1498 TTCTTGGTGGCTTGTCTGGGAGAGGACTTCATTTCCAAAGGATTCGCCACCGGATTC 1557 Qy  
1270 TTCTTGGAAGCGGTCTGGACGACCGCGATTTTCAGGCTGGGCGCATCACCACTCGTTTC 1329 Db  
1558 ATTGCGGATCACCGCGACCTTCCTTCAAGGCTCCACTGCTGATGATGAGCAGGACGCAATC 1617 Qy  
1330 ATCGAAGAGCGCGCGATTTGCTCACCGCGCGAGTTGCGCGACCGCGGACCAAGATC 1389 Db  
1618 CTGATTAATTGGAGATGTACCGTGAAACAGCTCATGTGTGCTGCTCAAGATGTT 1677 Qy  
1390 CTCAACTACCTGCGGACGCTGACCGTCAACAGCGGACCGGAGCGCCGCTCGCGGTG 1449 Db  
1678 GCAGCTCTATCATAGCTGCCTTAACATCAAGGATCTGCACTGCGACGCGCTTCCCGT 1737 Qy  
1450 TACCGCATGACAGCTGCCGACATCGACTTGGCGGCGAGCCGCGCGGCGCTCCAG 1509 Db  
1738 GACCGCTGAAGAGCTTGGCCGACCGCGGCTTGTCTGCTGATCTCGTGAGCAGGACGA 1797 Qy  
1510 CAGCGTTGACCGAATCTCGGCGCGGACGATTCGCGGCTTGGCTGCGGAGTTCGGCGGC 1569 Db  
1798 CTGCGAGTTACTGATACCACTTCGCGGATGACACCAAGCTTTTGTCTGCGACCCGAGTC 1857 Qy  
1570 GTGCGGCTACCGACACCACTTCGCGGACCGGACCAAGTGTGCTGGCGACCCGCGTC 1629 Db  
1858 CGCTCATTCGACCTGAAGAGCTGCGGAGAGGCGGTGCAAGCTGACTCTGAGCTTTTG 1917 Qy  
1630 CGCACAGCGGCTGCTCAAGGTGGCACTTACATCGCACGGAACCATTCGCGAGCTGTTG 1689 Db  
1918 TCCGTGAGGCTCGGCGCGGACCTACGATGTGGCGATGCTTTCCTTTGAGGAT 1977 Qy  
1690 TCGTGGAGTGTGGGCGGAGCGACTTACGCTGGCGCTGCGGTTTCTCAAGGAGGAC 1749 Db  
1978 CCGTGGACAGGCTCGAGCTGCGGAGCGGATGCCGAATGTAAATTCAGATGCTG 2037 Qy  
1750 CCGTGGAGCGGCTGACGGGCTCGCGGAGCGATGCCGAACATCTGCTTCAAGTGTG 1809 Db  
2038 CTTGCGCGCGCAACACCGTGGGATACACCCGTACCCAGACTCCGCTCCCGCGGTTT 2097 Qy  
1810 CTGCTGGGCTTAACACGTTGGGTACACGCTTATCCGAGAGCGGTACGACGCGCTTC 1869 Db  
2098 GTTAAGGAAGCTCGAGCTCGGCGTGACATCTTCGCAATTCGACGCGCTTAACGAC 2157 Qy  
1870 GTCGCGGAGGCGACCGGAGACCGGATCGACATCTTCGCGATCTTCGACGGTTTGAACAAC 1929 Db  
2158 GTCTCCAGATGCTCCAGCAATCGACGCTGCTGGAGCCACACCGCGGTAGCGGAG 2217 Qy  
1930 GTCGACTGATGCGCGCGGCGCATCGACGCGTTCGCGAAGCCGAGCGGCATAGCCGAA 1989 Db  
2218 GTGGCTATGGCTTATCTGCTGATCTCTCTGATCCAAATGAAAGCTCTTACACCTGGAT 2277 Qy  
1990 GTGGCGATGTCCTACACCGCGACCTGTCCGATCCGCGGAAAGCTTTTACACGCTGGAC 2049 Db  
2278 TACTACCTAAAGATGGAGAGGATGTGCAAGTCTGCGCTACATCTTGGCCATTAAG 2337 Qy  
2050 TATTACCTGAAGCTGGCGGACCAAGATCGTCAAGCGGCGGCGACCTGCTGGCGATCAAG 2109 Db

2338 GATATGCTGGTCTGCTTGGCCAGCTGGGTAAACCAAGCTGATCACCGCACTGGCCCT 2397 Qy  
2110 GACATGGCGCGCTGTGCGCGCACCGCGCGCGACGCTGTGTGCGGCTGAAGTGG 2169 Db  
2398 GAATTCGATCTGCGACGTGACCGTGCACCCACGACACTCGGCTGCGGCTGGCAACC 2457 Qy  
2170 CGGTTGACCTGCGGTGACGCTGCACCCCATGACACCGCGGCGGACAGCTGGCGAC 2229 Db  
2458 TACTTTGCTCAGCTCAAGCTGGTGAGATGCTGTTGACGCTGCTTCCGACACCACTGTT 2517 Qy  
2230 TACGTGCGGCTCGGACGCGCGGCGGATGCGTGCACGCGCGCGCTCGCC 2289 Db  
2518 GGCACCACTCCAGCAATCCCTGTCTGCCATTTGTTGCTGCACTTCGGGACACCCCTCGC 2577 Qy  
2290 GGAACCAACGACGCGGCACTATCGTGCATGCTGGCGCGCGGCGGCAACAGCGAGTAT 2349 Db  
2578 GATAACCGTTTGAAGCTCGAGGCTGTTTCTGACCTCGAGCCGCTACTGGGAAGCAGTGGC 2637 Qy  
2350 GACACCGGTTGCTCCCTGCGCGGCTGCGATCTGAGCCGCTACTGGAGGCGCTGCGA 2409 Db  
2638 GGAATGCTACCTGCAATTTGAGTCTGAAACCCAGGCGGCAACCGGTGCTTACCGCCAC 2697 Qy  
2410 AAGGTATACCTCCCTTCGAATCCGCTTTCGCGCGCGGACCGCGGCTGATACCCAC 2469 Db  
2698 GAAATCCAGCGGACAGTTGTCCAACTGCTGTCACAGGCCACCGCACTGGGCTTGGC 2757 Qy  
2470 GAGATCCCGGCGGCGCAATTTGTGCAATCTGCGTCAGCAGGCGATTTGGCTGGGACTCGC 2529 Db  
2758 GATCGTTTGAATCTCATGGAAGCACTAGCGAGCGTTAATGAGATGCTGGGACGCGCA 2817 Qy  
2530 GACCGTTTCGAGGACATCGAAAGCGCTACCGCGCGCGGCGGCGATTTTGGGACACCTG 2589 Db  
2818 ACCAAGTCAACCCATCTCCAAAGTTTGGGACCTCGCACTCCACCTCGTTGGTGGC 2877 Qy  
2590 GTCAAGGTCAACCGTCCAGCAGGTGCTGGCGACCTGCGCTGGCCCTGTGGTGGC 2649 Db  
2878 GGTGTGATCCAGCAGACTTTGTGCGGATCCAAAGTACGACATCCGAGACTCTGTC 2937 Qy  
2650 GGTGTGCGCGCGGAGGACTTTCGCGAAGAGCCCTCCGCTACGACATCCCGGATTCGCTG 2709 Db  
2938 ATCCGTTTCTCGCGCGAGCTTGTAACTCCAGTGTGCTGGCGGCGAGGCACTGCGC 2997 Qy  
2710 ATCGCTTCTGCGCGGAGCTCGGTGACCCGCGCGGCTGGCGGAGCGCTTGGC 2769 Db  
2998 ACCCGCGCATGGAAGCGCGCTCCGAAGGAGGACCTCTGAGCGAAGTTCCTGAGGAA 3057 Qy  
2770 ACCAAGGCTTTACAGGCGCGGCGCGGCGGCGGCGGAGCGCGCTGACGCGCGGAGAC 2829 Db  
3058 GAGCAGGCGCACCTCGAGGCTGATGATTCCAAAGAACGTGCAATAGCTCAACCGGCTG 3117 Qy  
2830 GAAGCGGCTTGGCGCGCGCGCGGCTG-----CGCGACGGCAGGCGGCTGAACCGGCTG 2883 Db  
3118 CTGTTCCCGAAGCAACCGAAGAGTTCTTGAGACACCGTCCGCTTCGGCAACACCTCT 3177 Qy  
2884 CTGTTCCCGGACCGACCAAGAGCTCGAGGACACCGCGAGAGTACGCTGACACCTCG 2943 Db  
3178 GCGCTGATGATGCTGATTTCTTACGGCTGTTCGAAGGCGCGGAGACACCGGCTGAA 3237 Qy  
2944 GGGCTGAGCGGCGAACCAAGTTCTTACGATTTGGGAGGGCGGAGACACCGGCTGAA 3003 Db  
3238 CTGCGAGATGTGCGCACCCCACTGCTTGTGCTGATGCGATCTCTGAGCGAGAGAT 3297 Qy  
3004 CT---CGAGCGGCGGTGAGTTGCTGATCGGCTGGAGGCCATCTCCGATCCCGACGA 3060 Db  
3298 AAGGTATGCGCAATTTGTGCGCAACGTCAGCGGCGGATCCGCGCAATCGGTGCGGT 3357 Qy  
3061 CGCGCATGCGCACCGTGTGCTTCTCAACGGGCGAGCTGCGGCGGCTGCTGCTGCGC 3120 Db  
3358 GACCGCTCGCTGATCTGTCAACGCAACCGCAGAAAGGCGAGATTCCTCAACAGGCGC 3417 Qy  
3121 GACCGAGCATCGCCATCGAGTGTGCGCGCGCGGAGAGGCGGACCGGCGCAACCGGAC 3180 Db



730 GCGCGCAGCTGACGCGCGAGTTGGTTACAGAGATGTGCGTCGATGCGGTGCGCTTCGCC 789  
1018 CGCTCCATTGGTTACGAGCGCGCGGGAACCGTGGGAATTTGGTTCGATGAAAGGGCAAC 1077  
790 CGCCATATCGGGTACAGCTGCGCGGGCACCGTGCAGTTCTGCTGGAGCGAGGGGAG 849  
1078 CAGCTCTTCATCGAATGAACCAACGATATCCAGGTTGAGACACCGTGCATCGAAGAGTC 1137  
850 TATGTCCTTTCATCGAGATGAATCGCGGGTTAGTGGAGCACACCGTGCAGGAGATT 909  
1138 ACCGAGGTGGAACCTGGTGAAGGGCGAGATGCGCTGGCTGGTGGTGCACCTTGAAGAA 1197  
910 ACCGAGCTCGACCTGGTCCAGCCAGCTGCGCATTCGCGCGGGAGACGCTCGAACAA 969  
1198 TTGGGTCTGACCAAGATGAAGATCAAGACCCAGGTGACGACTGCGAGTGCAGCTACCC 1257  
970 TTGGGCTGCGGCGAGAGACATCGCACCGCATGGCGCGCATACAGTGCAGCTACCC 1029  
1258 ACCGAGATCCAAACAAACCGCTTCGCGCCAGATACCGGAATATCACCGGTGACCGTCA 1317  
1030 ACCGAGATCCGCGCAACCGCTTCGCGCGCGGACACCGGCGCGATCAGCGCGTTGCGCAC 1089  
1318 CAGGCGGAGCTGCGCTGCTGACGTTGACGTTGACGTTGACGTTGACGTTGACGTTGAC 1377  
1090 GCGCGGCTGCGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 1149  
1378 CACTTTGACTCCATGCTGTTGAAATGACCTGCGGTGTTGCGACTTTGAACTGCTGTT 1437  
1150 TACTTGCATCCATGCTGTTGAACTGACCTGCTGCGGCGCGTGCATCCCTTACCGCAGTG 1209  
1438 GCTGTTGCAAGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1497  
1210 AGCGCTGCGCGCGCGGAGTTCGCGGAGTTCCGATCCGCGGCGGATGCGAGATATTCG 1269  
1498 TTCTTGGTCTGCTGCTGCGGAGAGACTTCACTTCCAAAGCGCATGCCACCGGATTC 1557  
1270 TTCTGCAAGCGGTTCTGATGACCGCGACTTCCGAGCGGCGCGGCTCACACGCTCTTC 1329  
1558 ATTGCGCATACCGGCTCTCTTCCAGGTTCCAGCTGCTGATGATGAGCAGGAGCGCATC 1617  
1330 ATTGATGAGCGCGCGAGTCTGACGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCGCG 1389  
1618 CTGATTTACTTGGCAGATGTCAACGTTGAAACAGCTCATGTTGCTGCTGCTGCTGCTGCT 1677  
1390 CTAACTTCTGCGCGATGTCAACGTTGAAACAGCTCATGTTGCTGCTGCTGCTGCTGCT 1449  
1678 GCGCTCCTATCGATAGCTGCTTCAATCAAGATGCTGCGCTGCGCGCGCGCTGCGCGCG 1737  
1450 TACCGGACGACAGCTGCGCGCGCTTCTGATCTGCGGCGCGCGCGCGCGCGCGCGCGCG 1509  
1738 GACCGCTGAGAGAGCTTGGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1797  
1510 CAGCGCTAGTCAAGTTGGCGCGCGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1569  
1798 CTGCGATTTACTGATACACCTTCCGCGATGACACCGAGTCTTGGTGGAGCGCGCGCTC 1857  
1570 GTCGGGTACCCGATACACATTCGCGGATGCTCACAGTCTGTTGCTGCTGCTGCTGCTG 1629  
1858 CGCTCATTTGCGCATGAAAGCTGCGCGCAGAGCGCTGCGAAAGCTGATCTCTGAGCTTTTG 1917  
1630 CGACACAGCGACTGTGCGGGTGGCACCGTATCTGCGCGGACCATGCGCGAGCTGTTG 1689  
1918 TCCGTGAGGCGCTGGCGCGCGCGACCTACATGATGCGGATGCGGTTCTTCTTTGAGGAT 1977  
1690 TCCGTGAGGCTTGGCGCGCGTGCAGCTTACGATGCGCGCTGCGCTTCTTCAAGGAGGAT 1749  
1978 CCGTGGGACAGCTCGAGCTGCGCGCGCGATGCGGATGCGGATGCGGATGCGGATGCGGAT 2037  
1750 CCTTGGGACAGCTGCGGCGCGCACCTGCGTGGAGCATGCGGATGCGGATGCGGATGCG 1809  
2038 CTTGCGCGCGCGACACCGTGGGATACCCCGGTACCGAGCTCCGCTGCTGCGCGCGGTTT 2097  
1810 TTGCGGCGCGCGCAATACCGTGGGCTACACCGCGTACCGCGGAAATCGTGAGCTGCGGCTT 1869

2098 GTTAAGGAAGCTGCCAGCTCCGCGTGGACATCTTCCGCATCTTGGACGCGCTTAACGAC 2157  
1870 GTCCAGAGAGCAACGCGCACCGGTATCGACATCTTGGTATCTTGGACGCGCTCAACAC 1929  
2158 GTCTCCAGATGCGTCCAGCAATCGAGCGAGTCTTGGAGACCAACACCGCGGTAGCGGAG 2217  
1930 ATCGAGTTCGATGCGTCCGCGATCGAGCGAGTACGCGAAACAGGTTCTCGGATAGAGAA 1989  
2218 GTGGCTATGGCTTATTTCTGCTGATCTCTCTGATCCAAATGAAAGCTTACACCTCGGAT 2277  
1990 GTCCGATGCTTACACGCGCGGACCTTACCGATCCGGGTGAACAGCTGTACACGCTGGAC 2049  
2278 TACTACTAAAGATGGAGAGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAG 2337  
2050 TACTACTGAAACTGGCTGAGCAGATCGTGGAGCGCGCGCCCATGTGCTGGGATCAAG 2109  
2338 GATATGCTGCTGCTGCTGCGCCAGCTGGGTAAACCAAGCTGTGTACCCGCACTGCGCCGT 2397  
2110 GATATGCTGAGATGTTGGCGCGCGCGCTCAAGGTTGGTCAAGCGCCCTGCGCAGT 2169  
2398 GAATTCGATGCGAGTGCACGCTGCACACCGAGACACTGCGGCTGGCGAGTGGCAAC 2457  
2170 CGCTTCGACCTGCGCTTCACTGTCACACCGAGCACACCGGCTGGCGAGTTCGCGAC 2229  
2458 TACTTTGCTGAGCTCAAGCTGCTGCGATGCTGTTGACGCTGCTTCCGACCACTGTCT 2517  
2230 TATGTCGCGCTTGGACGCGCGGCGGATGCGCTGCGAGCGCGCGCGCGCTGGCG 2289  
2518 GGCACCACTCCAGCGCATCCCTGCTGCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2577  
2290 GMAACGACCAACGCGCGCTGAGCTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2349  
2578 GATACCGGTTGAGCTCGAGGCTGTTTCTGACTCGAGCGCTACTGGGAGAGCAGTGGCG 2637  
2350 GACACCGCTGCTGCTTTCGCGGCTGCTGCGCTGCGAGCGCTACTGGAGGCGTTACGA 2409  
2638 GGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2697  
2410 AAAGTGTATGCGCGTTTCGAGTCTGGGTTGCGGCGCGGCGGCGGCTTATCACAC 2469  
2698 GAAATCCAGCGGAGAGTTCGCAACCTGCGTGCACAGGCGCGCGCGCTGCTGCTGCTGCT 2757  
2470 GAGATTCGCGCGCGCGCACTGCTCAATCTGCGCGAGCAAGCAATGCTGCTGCTGCTGCT 2829  
2758 GATGCTTTCGAACTCATCGAGAGCACTACGAGCGCGTTTAAATGAGATGCTGGAGCGCCA 2817  
2530 GATCGATTCGAGAGATCGAAGAGGCTACGCGCGCGCGCGAGTGTGGGCGAGCTG 2589  
2818 ACCAGGTCAACCATCTCTCGAGGTTGTTGGGAGCTTCCGACTCCACCTGCTGCTGCTGCT 2877  
2590 GTTAAGGTTCAGCGCAACGTCGAGGTTGCTGGGAGCTGCTGCGCTGCGGCTGCGGCTGCT 2649  
2878 GGTGATGATCAGCAGACTTTGCTGCGGATCCCAAAAGTACGACATCCGAGACTCTGCT 2937  
2650 GGTGTCAGTCAGACGATTCGCTCGATCCGATCCAGCGGATTTGGCATCCCGGAATCGTA 2709  
2938 ATCGGTTCTTGGCGGAGCTTGGTAAACCTTCCAGTGGCTGGCGAGCGGCTGCTGCTGCT 2997  
2710 CTGCGATTTCTGCGGCGGAGCTGCGGATCCGCGCGCGGCTGGCGGAGCGGCTGCTGCT 2769  
2998 ACCCGCGCATTCGAGGCGGCTCCGAGGCGAGGCACTCTGACGGAAGTTCTCTGAGGAA 3057  
2770 ACTGCGGCGCTGCGCGGCTGCGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2826  
3058 GAGCAGGCGCACTCGAGCTGATGATTCGAGGAGAGTTCGCAATAGCTCAACCGGCTG 3117  
2827 GACGAGATGCGCTTATCGTGGTGGAGCGCAAG---CGTACGCGCGCGGCGGCGGCGGCTG 2883  
3118 CTGTTCCGAGGCGCAACCGGAGAGTTCCTCGAGCACCGTTCGCGCTTCCGCGCAACCTCT 3177  
2884 TTATTTCCAGTCCAAAGAAAGGATTCATGAGCACCGGGAAGCTTACGCGGACACGCTCG 2943

QY	3178	GGCTGGATGATCGTGAATCTTCTACGGCTGGTCTGAAGCCGGGAGACTTTGATCCG	3237
Db	2944	CAATTGTGGCCCAACGAGTCTTCTATGCTTCGCCCAAGGTGAAGCATCGGTGAAG	3003
QY	3238	CTGCCAGATGTGCGCACCCCACTGCTGTTTCGCCCTGGATGCGATCTCTGAACGACGAT	3297
Db	3004	CTG---GAGCGTGGGTGAGCTGTTGATCGGGCTGGAGGCCATTTCCGAACCCGACGAA	3060
QY	3298	AAGGTATCGCAATGTTGTGCGCAACGTCAACGCGCGAGATCCGCCCCAATCGTGTGGT	3357
Db	3061	CGCGGATCGCAACGATGATGTCATCTCAACGGCGAGCTGCGGCCGCTAGTGGC	3120
QY	3358	GACCGCTCGGTGAGTCTCTACCCGCAACCGCAGAAAAGGCGAGATTCCTCAACAGGGC	3417
Db	3121	GACCGCAGATTTGCCAGTCCGCTTCGCGCGCGCGAGAGGCCGACCGCGCAATCCCGGA	3180
QY	3418	CATGTTGCTGCAACCATTCGCTGGTGTTCACCGTGAATGTTGCTGAAGGTGATGATGC	3477
Db	3181	CACATCGCGCGCCATTTTCGCGAGTCGTCACCGTTCGCGGTGCGGTGCGGTGCGGTG	3240
QY	3478	AAGGCTGGAGATGCACTCGCAATCATCGAGGCTATGAAGATGAAGCAACATCACTGCT	3537
Db	3241	GGCGCGCGCAACCATCGCCACCATCGAGCGGATGAAGTGAAGCCCGCATCCCGCC	3300
QY	3538	TCGTTGACGGCAAAATCGATCGCTGTTGCTGCTGCTGCAACGAGGTGAAGGTGCG	3597
Db	3301	CCGTTTGCAGCGCACCTGAGCGGGTGGCGGTGTCGGACACCGCCGAGGTGAGGGCGGA	3360
QY	3598	GACTTCATGCTGCTGCTTCTTAA 3621	
Db	3361	GACCTTGTGCTGTGAGCTGA 3384	

RESULT 15  
 AAI99682\_33/c  
 Continuation (34 of 45) of AAI99682 from base 3300001 (Mycobacterium tuberculosis strain  
 WP Sequence split into 45 fragments LOCUS AAI99682 Accession Aai99682

Fragment Name	Begin	End
AAI99682_00	100001	110000
AAI99682_01	200001	210000
AAI99682_02	300001	310000
AAI99682_03	400001	410000
AAI99682_04	500001	510000
AAI99682_05	600001	610000
AAI99682_06	700001	710000
AAI99682_07	800001	810000
AAI99682_08	900001	910000
AAI99682_09	1000001	1010000
AAI99682_10	1100001	1110000
AAI99682_11	1200001	1210000
AAI99682_12	1300001	1310000
AAI99682_13	1400001	1410000
AAI99682_14	1500001	1510000
AAI99682_15	1600001	1610000
AAI99682_16	1700001	1710000
AAI99682_17	1800001	1810000
AAI99682_18	1900001	1910000
AAI99682_19	2000001	2010000
AAI99682_20	2100001	2110000
AAI99682_21	2200001	2210000
AAI99682_22	2300001	2310000
AAI99682_23	2400001	2410000
AAI99682_24	2500001	2510000
AAI99682_25	2600001	2610000
AAI99682_26	2700001	2710000
AAI99682_27	2800001	2810000
AAI99682_28	2900001	2910000
AAI99682_29	3000001	3010000
AAI99682_30	3100001	3110000
AAI99682_31	3200001	3210000
AAI99682_32	3300001	3310000
AAI99682_33	3400001	3410000
AAI99682_34	3500001	3510000

WP	AAI99682_35	3500001	3610000
WP	AAI99682_36	3600001	3710000
WP	AAI99682_37	3700001	3810000
WP	AAI99682_38	3800001	3910000
WP	AAI99682_39	3900001	4010000
WP	AAI99682_40	4000001	4110000
WP	AAI99682_41	4100001	4210000
WP	AAI99682_42	4200001	4310000
WP	AAI99682_43	4300001	4410000
WP	AAI99682_44	4400001	4411529

Query Match 35.7%; Score 1294.4; DB 4; Length 110000;  
 Best Local Similarity 62.0%; Pred. No. 0;  
 Matches 2099; Conservative 0; Mismatches 1276; Indels 9; Gaps 3;

QY	238	AAGATCTTGTAGCAACCGCGCGGAAATCGCGTCCGTGCTTTCCGTGCGAGCACTCGAA	297
Db	23034	AAGTGTCTGTGCCAATCGCGGGAGATCGGATCCGGGCCCTTTCTGTGCGGCTACGAA	22975
QY	298	ACCGGTGACGACCGTAGTATTACCCCGGTGAAGATCGGGATCATTCACCGCTCT	357
Db	22974	CTGGGCTCGGAACCGTGGCCGTTTATCGGTACGAGGACCGCAATTCGCGACCGCTCTC	22915
QY	358	TTTGTCTTCAAGCTGTCCGCAATGTACCGAAGGTCAACAGTCAAGGCGTCAAGCGTACCTGGAC	417
Db	22914	AAGCGGACGAGTCTTACAGATCGCGGACATCGGTCAACCGGTGATCATACCTGTG	22855
QY	418	ATCGATGAATATCGGTGCGCTAAAGTTAAAGAGATCCCATTTTACCGGATAC	477
Db	22854	GTGACGAGATCGTCGCGACGCGCGTTCGGGCGGGTGCAGACGCTATCTACCTGCTAC	22795
QY	478	GGCTTCCTGCTCAAAATGCCAGCTTGGCCGCGAGTGTGCGGAAAACGGCATATCTTTT	537
Db	22794	GGTTCCTATCGGAGATTCGGATCTGGCTGCGCATGCGGCGGGGCGATCAGCTTC	22735
QY	538	ATTGGCCCAACCCAGAGTCTTGTATCTCACCGGTGATAAGTCTCGCGGGTAAACGCC	597
Db	22734	GTGCTCCAGCGCGAAGTCTTGTAGTGGCTGGGAATAAGTCTCGCGCATCGCGCG	22675
QY	598	GCGAAGAGGCTGTCTGCCAGTTTGGCGGAATCACCCCGAGCAAAACATCGATGAG	657
Db	22674	GCCGCGAAGCGGCTTGGCGGCTGTATGTCTCGGCGCGCTCGGCTCGGTGACGAA	22615
QY	658	ATCGTTAAAGCGCTGAAGCGCAGACTTACCCCATCTTTGTGAAGCAGTTGCGGFGT	717
Db	22614	CTGCTGCGTTGGCGCGCATGCCGTTTCTCGGTTCGCTCAAGGCGATGCGGTTGC	22555
QY	718	GGCGACGCGGTATGCGTTTGTGCTTCACTGATGAGCTTCGCAATTAGCAACGAA	777
Db	22554	GGGCGCGGGGTATGCGTGTGTCGCGCATATCGCGCGCTTCGAGGCGATCGAAGCC	22495
QY	778	GCATCTGTGAAGCTGAAGCGGCTTTCGCGGATGCGCGGCTATATGTGCAACGCTGTG	837
Db	22494	GCGAGCGGAGCGGATCGGCTTTCGCGGACCCGACGCTTCTATCTCGAGCAGCAGTG	22435
QY	838	ATTAACCTCAGCATATTGAAGTGCAGATCTTGGCGATCACACTGAGAGTTGTACAC	897
Db	22434	ATCAATCCACCCCATCGAGGTGAGATTCGCGGCAACCTTCGCGGCGATCGATCAT	22375
QY	898	CTTTATGAACGTGACTGCTCACTCAGCGTGTCCACCAAAAAGTTGCGAAATTCGCGCA	957
Db	22374	CTCTATGAGCGTGTGAGTGTGCGGTGCGCATCAGAGGTATCATCGAGCTCGCGCC	22315
QY	958	GCACAGCATTTGGATCCAGAACTCGCTGATCGCATTTGTGCGGATGCAATTAAGTTCTGC	1017
Db	22314	GCGCGCACCTGGACCGCGAGTTCGTTTACAAGATGTGCGTGTGCGGTGCGCTTGC	22255
QY	1018	CGCTCCATTGGTTACGAGGCGCGGAACCGTGAATTTCTGGTCTGATGAAAGGCAAC	1077
Db	22254	CGCATATCGGATACAGCTGCGCGGACCGCTGAGTTCTGCTGACGAGGGGAG	22195
QY	1078	CACGCTTTCATCGAAATGAACCCGATTCAGGTTGAGCACCGGTGACGACCGTGTG	1137





Search completed: March 23, 2004, 18:35:43  
Job time : 1318 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 17:20:33 ; Search time 237 Seconds  
(without alignments)  
8478.807 Million cell updates/sec

Title: US-10-045-072-1  
Perfect score: 3621  
Sequence: 1 tggggcgggtagtagctctg.....tgatcgctgctgttctctaa 3621

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgm2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgm2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgm2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgm2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgm2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*  
6: /cgm2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3621	100.0	3621	3	US-09-220-081-1
2	3621	100.0	3621	4	US-09-677-575-1
3	1294.4	35.7	4403765	3	US-09-103-840A-2
4	1294.4	35.7	4411529	3	US-09-103-840A-1
5	1249.6	34.5	32155	4	US-08-311-731A-1
6	509.2	14.1	3492	4	US-09-134-000C-2302
7	486.2	13.4	3465	4	US-09-134-001C-591
8	479.8	13.3	5030	4	US-08-956-171E-324
9	476.2	13.2	2922	4	US-09-107-532A-1156
10	313.4	8.7	1362	4	US-09-543-681A-1699
11	313	8.6	1362	1	US-07-956-700B-5
12	313	8.6	1362	1	US-08-476-537-5
13	313	8.6	1362	1	US-08-485-607-5
14	313	8.6	1362	2	US-08-475-879-5
15	313	8.6	1362	1	US-09-433-043B-5
16	311.4	8.6	1362	1	US-08-611-107-7
17	311.4	8.6	1362	2	US-08-422-560A-7
18	311.4	8.6	1362	3	US-08-468-793-7
19	299	8.3	3077	1	US-08-074-121-1
20	299	8.3	3077	5	PCT-US94-06447-1
21	298.6	8.2	3065	1	US-07-956-700B-1
22	298.6	8.2	3065	1	US-08-476-537-1
23	298.6	8.2	3065	1	US-08-485-607-1
24	298.6	8.2	3065	1	US-08-611-107-5
25	298.6	8.2	3065	2	US-08-422-560A-5
26	298.6	8.2	3065	2	US-08-475-879-1
27	298.6	8.2	3065	3	US-08-468-793-5

28	298.6	8.2	3065	4	US-09-433-043B-1
29	297.4	8.2	2907	4	US-08-252-991A-2354
30	288.6	8.0	1511	4	US-09-634-238-58
31	284.8	7.9	1425	4	US-09-328-352-3436
32	270.6	7.5	1921	3	US-08-662-344-1
33	263.8	7.3	1953	4	US-09-252-991A-2481
34	255.4	7.1	1664976	4	US-08-916-421B-1
35	251.6	6.9	1356	4	US-09-540-236-1127
36	251.6	6.9	119211	4	US-09-773-816-1
37	251.6	6.9	4403765	3	US-09-103-840A-2
38	250.2	6.9	4411529	3	US-09-103-840A-1
39	250.2	6.9	2031	4	US-09-252-991A-9572
40	249	6.9	3006	4	US-09-252-991A-9720
41	249	6.9	13702	4	US-08-961-537-7
42	247.4	6.8	1398	4	US-09-252-991A-10409
43	247	6.8	1446	4	US-09-252-991A-10752
44	246.8	6.8	1830121	4	US-09-557-884-1
45					

ALIGNMENTS

RESULT 1  
US-09-220-081-1  
; Sequence 1, Application US/09220081  
; Patent No. 6171833  
; GENERAL INFORMATION:  
; APPLICANT: Sinskey, Anthony J.  
; APPLICANT: Lessard, Philip A.  
; APPLICANT: Wallis, Laura B.  
; APPLICANT: Stephanopoulos, Gregory  
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum  
; FILE REFERENCE: 1533.079000  
; CURRENT APPLICATION NUMBER: US/09/220,081  
; CURRENT FILING DATE: 1998-12-23  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3621  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (139)..(3621)  
US-09-220-081-1

Query Match	100.0%;	Score 3621;	DB 3;	Length 3621;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3621;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	TGGGGCGGGTTAGATCCTGGGGGGTTTATTTTCATTACATTTGGCTTGAAGTCGTGCAGG	60	
Db	1	TGGGGCGGGTTAGATCCTGGGGGGTTTATTTTCATTACATTTGGCTTGAAGTCGTGCAGG	60	
Qy	61	TCAGGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAAAACGATGTTGATTCGGGG	120	
Db	61	TCAGGGGAGTGTTCGCCGAAACATTGAGAGGAAACAAAAACGATGTTGATTCGGGG	120	
Qy	121	AATCGGGGTTAGATCCTAGGCGGACGACGCTATACCCCTTGGCGGTCTCTTTGTTG	180	
Db	121	AATCGGGGTTAGATCCTAGGCGGACGACGCTATACCCCTTGGCGGTCTCTTTGTTG	180	
Qy	181	AAAGGAATTAATTAATCTAGTGTGCTACATCAATCTTCAACGCTTCAGCATTCAAAAG	240	
Db	181	AAAGGAATTAATTAATCTAGTGTGCTACATCAATCTTCAACGCTTCAGCATTCAAAAG	240	
Qy	241	ATCTTGTAGAAACCGCGGAAATCGCGGTTCGCTTCCGTGCGACACTCGAAACC	300	
Db	241	ATCTTGTAGAAACCGCGGAAATCGCGGTTCGCTTCCGTGCGACACTCGAAACC	300	
Qy	301	GGTCAGCCAGGTAGTCTATTACCCCGGTGAAGATCGGGGATCATTCACCGCTCTTTT	360	

Db 301 GGTGAGCCACGGTAGCTATTTACCCCGCTGAAGATCGGGGATCAATCCACCGCTCTTTT 360  
Qy 361 GCTTCTGAAGCTGTCCGCAATGTGTACCGAAGGCTCACAGTCAAGCGGTACCTGGACATC 420  
Db 361 GCTTCTGAAGCTGTCCGCAATGTGTACCGAAGGCTCACAGTCAAGCGGTACCTGGACATC 420  
Qy 421 GATGAATATTCGGTGCAGTCAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480  
Db 421 GATGAATATTCGGTGCAGTCAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480  
Qy 481 TTCTGTCTGTAATGCCCCAGCTTGGCCGCGAGTGTGCGGAAACCGCATATCTTTAT 540  
Db 481 TTCTGTCTGTAATGCCCCAGCTTGGCCGCGAGTGTGCGGAAACCGCATATCTTTAT 540  
Qy 541 GGCCCAACCCAGAGTTCTTGATCTCACCGGTGATAAGTCTCGCGGTAAACCCCGG 600  
Db 541 GGCCCAACCCAGAGTTCTTGATCTCACCGGTGATAAGTCTCGCGGTAAACCCCGG 600  
Qy 601 AAGAAGGCTGTCTGCCAGTTTGGCGGAATCCACCCCGAGCAAAACATCGATGATC 660  
Db 601 AAGAAGGCTGTCTGCCAGTTTGGCGGAATCCACCCCGAGCAAAACATCGATGATC 660  
Qy 661 GTTAAAGCCGTGAAGCCAGACTTACCCCATCTTTGTGAAGGCACTTCCCGGTGGTGC 720  
Db 661 GTTAAAGCCGTGAAGCCAGACTTACCCCATCTTTGTGAAGGCACTTCCCGGTGGTGC 720  
Qy 721 GGACGCGGTATGCGTTTGTGTCTTCACTGTAGCTTCGCAAAATTAGCAACAGAGCA 780  
Db 721 GGACGCGGTATGCGTTTGTGTCTTCACTGTAGCTTCGCAAAATTAGCAACAGAGCA 780  
Qy 781 TCTCGTGAAGCTGAAGCGGCTTTCGGGATGCGCGGTATATGCGAAAGTGTGAT 840  
Db 781 TCTCGTGAAGCTGAAGCGGCTTTCGGGATGCGCGGTATATGCGAAAGTGTGAT 840  
Qy 841 AACCTCAGCATATTGAAGTGCAATCCTTTGGCGATCACACTGAGAGTGTGACACCT 900  
Db 841 AACCTCAGCATATTGAAGTGCAATCCTTTGGCGATCACACTGAGAGTGTGACACCT 900  
Qy 901 TATGAACGTGACTGCTCACTGACGCTGTCGACCAAAAGTTGCGAAATGCGCGAGCA 960  
Db 901 TATGAACGTGACTGCTCACTGACGCTGTCGACCAAAAGTTGCGAAATGCGCGAGCA 960  
Qy 961 CAGATTTGGAATCAGAACTCCGTGATCGCAATTTGTGCGATGCACTGAAAGTCTGCGC 1020  
Db 961 CAGATTTGGAATCAGAACTCCGTGATCGCAATTTGTGCGATGCACTGAAAGTCTGCGC 1020  
Qy 1021 TCCATTTGTTACCGGGCGCGGAAACCGTGGAAATTTCTTGGTGAATGAAAGGGCAACC 1080  
Db 1021 TCCATTTGTTACCGGGCGCGGAAACCGTGGAAATTTCTTGGTGAATGAAAGGGCAACC 1080  
Qy 1081 GTCTTCATCGAAATGAACCCAGCTATCCAGTTGAGCACACCGTGAAGTCAAC 1140  
Db 1081 GTCTTCATCGAAATGAACCCAGCTATCCAGTTGAGCACACCGTGAAGTCAAC 1140  
Qy 1141 GAGTGGACCTGGTGAAGGCGAGATGCGTGTGCTGCTGAGCACTGAGAGGAAATG 1200  
Db 1141 GAGTGGACCTGGTGAAGGCGAGATGCGTGTGCTGCTGAGCACTGAGAGGAAATG 1200  
Qy 1201 GGTCTGACCCAGATPAAGATCAAGACCCAGCTGAGCACTGAGTGCAGTCAACACG 1260  
Db 1201 GGTCTGACCCAGATPAAGATCAAGACCCAGCTGAGCACTGAGTGCAGTCAACACG 1260  
Qy 1261 GAAGATCCAAACACCGGTTTCGGCCAGATACCGGAATATCAACCGGTACCGTCAAC 1320  
Db 1261 GAAGATCCAAACACCGGTTTCGGCCAGATACCGGAATATCAACCGGTACCGTCAAC 1320  
Qy 1321 GCGGAGCTGGGTTCTGCTTGAAGGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1380  
Db 1321 GCGGAGCTGGGTTCTGCTTGAAGGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1380  
Qy 1381 TTTGATCCATGCTGGTGAATAATGACCTCCCGGTTCGCACTTTGAAACTGCTGTGCT 1440  
Db 1381 TTTGATCCATGCTGGTGAATAATGACCTCCCGGTTCGCACTTTGAAACTGCTGTGCT 1440

Qy 1441 CBTGCAAGCGCGCTTGGCTGAGTTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
Db 1441 CBTGCAAGCGCGCTTGGCTGAGTTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
Qy 1501 TTGCGTGTGCTTGGTGGGAAAGGACTTCACTTCAAGCGCATCGCCACCGGATTCATT 1560  
Db 1501 TTGCGTGTGCTTGGTGGGAAAGGACTTCACTTCAAGCGCATCGCCACCGGATTCATT 1560  
Qy 1561 GCGATCACCGCACCTCTTCAAGGCTCCACCTGCTGATGATGAGGAGGAGGAGGAGG 1620  
Db 1561 GCGATCACCGCACCTCTTCAAGGCTCCACCTGCTGATGATGAGGAGGAGGAGGAGG 1620  
Qy 1621 GATTACTTGGCAGATGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 1680  
Db 1621 GATTACTTGGCAGATGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 1680  
Qy 1681 GCTCTTATGATGATGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 1740  
Db 1681 GCTCTTATGATGATGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 1740  
Qy 1741 GCGCTGAAGCAGCTTGGCCAGCGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
Db 1741 GCGCTGAAGCAGCTTGGCCAGCGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
Qy 1801 GCGTCTGATGATGCTTCCGCGATGCAACAGTCTTTCGCTGCTGCTGCTGCTGCTGCTG 1860  
Db 1801 GCGTCTGATGATGCTTCCGCGATGCAACAGTCTTTCGCTGCTGCTGCTGCTGCTGCTG 1860  
Qy 1861 TCATTTCGCACTGAAAGCCTTGGCGGAGGCGCTTCCGCTGCTGCTGCTGCTGCTGCTG 1920  
Db 1861 TCATTTCGCACTGAAAGCCTTGGCGGAGGCGCTTCCGCTGCTGCTGCTGCTGCTGCTG 1920  
Qy 1921 GTGAGGCTTGGCGCGCGGCTTCCGCGATGCAACAGTCTTTCGCTGCTGCTGCTGCTGCTG 1980  
Db 1921 GTGAGGCTTGGCGCGCGGCTTCCGCGATGCAACAGTCTTTCGCTGCTGCTGCTGCTGCTG 1980  
Qy 1981 TGGGACGCTCGAGCTCGCGGAGGCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
Db 1981 TGGGACGCTCGAGCTCGCGGAGGCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
Qy 2041 CCGCGCGCGAACAACGCTGGGATACACCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
Db 2041 CCGCGCGCGAACAACGCTGGGATACACCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
Qy 2101 AAGGAAGCTGCGAGCTCCGCGGTGAGCATCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
Db 2101 AAGGAAGCTGCGAGCTCCGCGGTGAGCATCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
Qy 2161 TCCGAGATCGCTCCAGCAATCGACGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220  
Db 2161 TCCGAGATCGCTCCAGCAATCGACGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220  
Qy 2221 GCTATGGCTTATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280  
Db 2221 GCTATGGCTTATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280  
Qy 2281 TACCTAAAGATGGCAGAGAGATGCTCAAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340  
Db 2281 TACCTAAAGATGGCAGAGAGATGCTCAAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340  
Qy 2341 ATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
Db 2341 ATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
Qy 2401 TTCGATCTGCCAGTGCAGTGCACACCGAGCATCTGCGGTGGCGAGTGGCAACCTTAC 2460  
Db 2401 TTCGATCTGCCAGTGCAGTGCACACCGAGCATCTGCGGTGGCGAGTGGCAACCTTAC 2460  
Qy 2461 TTTGCTGCAAGCTCAAGCTGGTGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520  
Db 2461 TTTGCTGCAAGCTCAAGCTGGTGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520

QY	2521	ACACCTCCAGCCATCCCTGTCTGCCATTTGTTGCTGCAATTCGCGCACACCCGTCGCGAT	2580
DB	2521	ACCACCTCCAGCCATCCCTGTCTGCCATTTGTTGCTGCAATTCGCGCACACCCGTCGCGAT	2580
QY	2581	ACCGGTTTCAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGAAGCAGTGCGCGA	2640
DB	2581	ACCGGTTTCAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGAAGCAGTGCGCGA	2640
QY	2641	CTGTACTCTGCCATTTGAGTCTGAAACCCAGGCCCAACCGGTCGCGTCTAACGCCACGA	2700
DB	2641	CTGTACTCTGCCATTTGAGTCTGAAACCCAGGCCCAACCGGTCGCGTCTAACGCCACGA	2700
QY	2701	ATCCAGGCGGACAGTTGTCCAACTCGCTGCGTGCACAGGCCACCGCACTGGGCCCTTGGGAT	2760
DB	2701	ATCCAGGCGGACAGTTGTCCAACTCGCTGCGTGCACAGGCCACCGCACTGGGCCCTTGGGAT	2760
QY	2761	CGTTTTCGAACTCATTCGAAGACAACTACGCAAGCCGTTAATGAGATGCTGGGACGCCCAACC	2820
DB	2761	CGTTTTCGAACTCATTCGAAGACAACTACGCAAGCCGTTAATGAGATGCTGGGACGCCCAACC	2820
QY	2821	AAGGTCAACCCATCTCTCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGTTGGGGT	2880
DB	2821	AAGGTCAACCCATCTCTCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGTTGGGGT	2880
QY	2881	GTGGATCCAGCAGACTTTGTCTGCCATTCGCAACAAAGTACGACATCCAGACTCTGTCTATC	2940
DB	2881	GTGGATCCAGCAGACTTTGTCTGCCATTCGCAACAAAGTACGACATCCAGACTCTGTCTATC	2940
QY	2941	GGGTTCTCGCGCGGAGCTTGTGTAACCTTCAGGTGGCTGGCCAGACCACTCGCGCACCC	3000
DB	2941	GGGTTCTCGCGCGGAGCTTGTGTAACCTTCAGGTGGCTGGCCAGACCACTCGCGCACCC	3000
QY	3001	CGCGCACTGGAAGGCCCTCCGAGGCAAGSCACTCTGACGGAAAGTTCCTGAGGAAGAG	3060
DB	3001	CGCGCACTGGAAGGCCCTCCGAGGCAAGSCACTCTGACGGAAAGTTCCTGAGGAAGAG	3060
QY	3061	CAGCGCACTTCGACGCTGATGATTCAGAGAACGTCGCAATAGCTTCAACCGGCTCGTG	3120
DB	3061	CAGCGCACTTCGACGCTGATGATTCAGAGAACGTCGCAATAGCTTCAACCGGCTCGTG	3120
QY	3121	TTCCGGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCGCTTTCGSCAACACCTCTGCG	3180
DB	3121	TTCCGGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCGCTTTCGSCAACACCTCTGCG	3180
QY	3181	CTGGATGATCGTGAATTTCTTACGGCTTGGTCGAAGCCGCGAGACTTTGATCCGCGCTG	3240
DB	3181	CTGGATGATCGTGAATTTCTTACGGCTTGGTCGAAGCCGCGAGACTTTGATCCGCGCTG	3240
QY	3241	CCAGATGTGCGCACCCCACTGCTTTGTCGCTTGGATGCGATCTCTGAGCCAGAGGATAAG	3300
DB	3241	CCAGATGTGCGCACCCCACTGCTTTGTCGCTTGGATGCGATCTCTGAGCCAGAGGATAAG	3300
QY	3301	GGTATGCGCAATGTTGTGGCCAAACGTCACGGCCAGATCCGCCCAATGCGTGTGCGTGAC	3360
DB	3301	GGTATGCGCAATGTTGTGGCCAAACGTCACGGCCAGATCCGCCCAATGCGTGTGCGTGAC	3360
QY	3361	CGCTCCGTTGAGTCTGTCAACCGCAACCGCAGAAAAGGCAGATTTCTTCCAAACAGGGCCAT	3420
DB	3361	CGCTCCGTTGAGTCTGTCAACCGCAACCGCAGAAAAGGCAGATTTCTTCCAAACAGGGCCAT	3420
QY	3421	GTTCTGCACCAATTCGCTGTTGTTGTCAACCGTGACTGTTGTGAAGGTGATGAGGTCAAG	3480
DB	3421	GTTCTGCACCAATTCGCTGTTGTTGTGTCAACCGTGACTGTTGTGAAGGTGATGAGGTCAAG	3480
QY	3481	GCTGGAGATGAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT	3540
DB	3481	GCTGGAGATGAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT	3540
QY	3541	GTTGACGGCAAAATCGATCGGTTGTGTTCTGCTGTCACACGAGGTGGAAGGTGGCGAC	3600
DB	3541	GTTGACGGCAAAATCGATCGGTTGTGTTCTGCTGTCACACGAGGTGGAAGGTGGCGAC	3600
QY	3601	TTGATCGTCGCTGTTTCTCTAA 3621	

```

Db          3601 TTGATCGTCGTCTTCCTRAA 3621
|||||
|||
RESULT 2
US-09-677-575-1
; Sequence 1, Application US/09677575
; Patent No. 6403351
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; APPLICANT: Stephanopoulos, Gregory
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790000
; CURRENT APPLICATION NUMBER: US/09/677,575
; CURRENT FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 09/220,081
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 1
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (199)..(3621)
US-09-677-575-1

```

QY 601 AAGAGGCTGGTCTGCGCAGTTTGGCGAATCACCCCGAGCAAAAACATCGATGAGATC 660  
DB 601 AAGAGGCTGGTCTGCGCAGTTTGGCGAATCACCCCGAGCAAAAACATCGATGAGATC 660  
QY 661 GTTAAAGCCCTGAGGCGAGCTTACCCCATCTTTGTGAAGGCACTTCCCGTGGTGGC 720  
DB 661 GTTAAAGCCCTGAGGCGAGCTTACCCCATCTTTGTGAAGGCACTTCCCGTGGTGGC 720  
QY 721 GGAGCGGTATGCGTTTGTGTTCTTCACTGATGAGCTTCGCAAAATAGCAACAGAGCA 780  
DB 721 GGAGCGGTATGCGTTTGTGTTCTTCACTGATGAGCTTCGCAAAATAGCAACAGAGCA 780  
QY 781 TCTGCTGAAGCTGAGGCGCTTTCCGCGATGCGCGGTATATCGGAAGCTGTGATT 840  
DB 781 TCTGCTGAAGCTGAGGCGCTTTCCGCGATGCGCGGTATATCGGAAGCTGTGATT 840  
QY 841 AACCCCTCAGCATATTGAAGTGCAGATCCTTTGGCGATCACACTGGAGAAGTTGTACACCTT 900  
DB 841 AACCCCTCAGCATATTGAAGTGCAGATCCTTTGGCGATCACACTGGAGAAGTTGTACACCTT 900  
QY 901 TATGAAGCTGACTCTCACTGCGAGCTGTCACCAAAAAGTTGCGAAATGCGCCAGCA 960  
DB 901 TATGAAGCTGACTCTCACTGCGAGCTGTCACCAAAAAGTTGCGAAATGCGCCAGCA 960  
QY 961 CAGCAATTGATCCAGAACTCGGTGATCGCAATTTGTGCGGATCGAGTAAAGTTCTGCCG 1020  
DB 961 CAGCAATTGATCCAGAACTCGGTGATCGCAATTTGTGCGGATCGAGTAAAGTTCTGCCG 1020  
QY 1021 TCCATTTGTTACCGAGGCGGGAACCGTGGATTTCTGTCGATGAAAGGCGCAACCC 1080  
DB 1021 TCCATTTGTTACCGAGGCGGGAACCGTGGATTTCTGTCGATGAAAGGCGCAACCC 1080  
QY 1081 GTCTTCATCGAAATGAACCCAGCTATCCAGCTTGAGCACACCGTGTGCTGAAAGTCA 1140  
DB 1081 GTCTTCATCGAAATGAACCCAGCTATCCAGCTTGAGCACACCGTGTGCTGAAAGTCA 1140  
QY 1141 GAGTGGACCTGGTGAAGGCGCAGATGGCTTGGCTGGTGGCACTTTGAAGGAATTG 1200  
DB 1141 GAGTGGACCTGGTGAAGGCGCAGATGGCTTGGCTGGTGGCACTTTGAAGGAATTG 1200  
QY 1201 GGTCTGACCCAGATTAAGATCAAGACCCACCGGTGAGCACTGAGTCCCGCATCACCA 1260  
DB 1201 GGTCTGACCCAGATTAAGATCAAGACCCACCGGTGAGCACTGAGTCCCGCATCACCA 1260  
QY 1261 GAAGATCCAAACACCGCTTCCGCGCAGATACCGGAATATCACCGGTACCGTCA 1320  
DB 1261 GAAGATCCAAACACCGCTTCCGCGCAGATACCGGAATATCACCGGTACCGTCA 1320  
QY 1321 GCGGAGCTGGGTGCTTGTGAGCGGTGAGTCACTGAGTCCGCTGCGGAAATCACCGCAC 1380  
DB 1321 GCGGAGCTGGGTGCTTGTGAGCGGTGAGTCACTGAGTCCGCTGCGGAAATCACCGCAC 1380  
QY 1381 TTTGACTCCATCTCGTGAATAATGACTGCGGTGCTTCCGACTTTGAACTGCTGTTGCT 1440  
DB 1381 TTTGACTCCATCTCGTGAATAATGACTGCGGTGCTTCCGACTTTGAACTGCTGTTGCT 1440  
QY 1441 CGTGACACGCGCTTGGTGAAGTTACCGGTGCTGCTGTTGCAACCAACTTGGTTTC 1500  
DB 1441 CGTGACACGCGCTTGGTGAAGTTACCGGTGCTGCTGTTGCAACCAACTTGGTTTC 1500  
QY 1501 TTGCGTGGCTTCTCGGGAAGAGGACTTCACTTCCAGCGCATGCGCACCGGATTCATT 1560  
DB 1501 TTGCGTGGCTTCTCGGGAAGAGGACTTCACTTCCAGCGCATGCGCACCGGATTCATT 1560  
QY 1561 GCGGATCACCGCATCTCTTCCAGCTCCACTGCTGATGATGAGCAGGAGCGATCTCTG 1620  
DB 1561 GCGGATCACCGCATCTCTTCCAGCTCCACTGCTGATGATGAGCAGGAGCGATCTCTG 1620  
QY 1621 GATTACTTGGCAGATGTCAACCGTGAACAGCTCATGTTGCGTCCAAAGGATTTGCA 1680  
DB 1621 GATTACTTGGCAGATGTCAACCGTGAACAGCTCATGTTGCGTCCAAAGGATTTGCA 1680

QY 1681 GCTCCTATCGATAGCTGCCCTTAACATCAAGGATCTGCCACTGCGACGCGGTTCCTCGTGAC 1740  
DB 1681 GCTCCTATCGATAGCTGCCCTTAACATCAAGGATCTGCCACTGCGACGCGGTTCCTCGTGAC 1740  
QY 1741 CGCCTGAAGCAGCTTGGCCCGCGCGTTCCTCGTATCTCCGTGAGCAGGACGACTG 1800  
DB 1741 CGCCTGAAGCAGCTTGGCCCGCGCGTTCCTCGTATCTCCGTGAGCAGGACGACTG 1800  
QY 1801 CGAGTTACTGATACCACTTCGCGGATGACACACAGTCTTTGTTGCGACCCGAGTCCGC 1860  
DB 1801 CGAGTTACTGATACCACTTCGCGGATGACACACAGTCTTTGTTGCGACCCGAGTCCGC 1860  
QY 1861 TCATTTCGCACTGAAGCCTGCGGAGAGCCCTGCGAAAGCTGACTCTCGAGCTTTGTCC 1920  
DB 1861 TCATTTCGCACTGAAGCCTGCGGAGAGCCCTGCGAAAGCTGACTCTCGAGCTTTGTCC 1920  
QY 1921 GTGAGGCTTGGGCGGCGGACCTTACGATGTCGATGCTTTCCTTTTGGAGTCCG 1980  
DB 1921 GTGAGGCTTGGGCGGCGGACCTTACGATGTCGATGCTTTCCTTTTGGAGTCCG 1980  
QY 1981 TGGGACAGGCTCGAGAGCTGCGGAGCGGATGCGAAATGTAACATTCAGATGCTGCTT 2040  
DB 1981 TGGGACAGGCTCGAGAGCTGCGGAGCGGATGCGAAATGTAACATTCAGATGCTGCTT 2040  
QY 2041 CGCGCCGCAACACGCTGGGATACACCCCTTACCCAGCTCCGTCGCGCGCTTTGTT 2100  
DB 2041 CGCGCCGCAACACGCTGGGATACACCCCTTACCCAGCTCCGTCGCGCGCTTTGTT 2100  
QY 2101 AAGGAAGCTGCGAGCTCCGCGGTGAGCATCTTCGCGCATCTTCGACGCGCTTAAACGATC 2160  
DB 2101 AAGGAAGCTGCGAGCTCCGCGGTGAGCATCTTCGCGCATCTTCGACGCGCTTAAACGATC 2160  
QY 2161 TCCGAGTGGCTCCAGCAATCGACGAGCTCGGAGACCAACCGCGGTAGCGAGGTG 2220  
DB 2161 TCCGAGTGGCTCCAGCAATCGACGAGCTCGGAGACCAACCGCGGTAGCGAGGTG 2220  
QY 2221 GCTATGGCTTATTTCTGGTGTCTCTCTGATCCAAATGAAAAGCTCTACACCTCGATTAC 2280  
DB 2221 GCTATGGCTTATTTCTGGTGTCTCTCTGATCCAAATGAAAAGCTCTACACCTCGATTAC 2280  
QY 2281 TACCTTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAAGAT 2340  
DB 2281 TACCTTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAAGAT 2340  
QY 2341 ATGGCTGGTCTGCTTCCGCGAGCTGCGGTAAACAGCTGGTCAACCGCATCGCGGTGAA 2400  
DB 2341 ATGGCTGGTCTGCTTCCGCGAGCTGCGGTAAACAGCTGGTCAACCGCATCGCGGTGAA 2400  
QY 2401 TTGCAATTCGCCAGTGCAGTGCACACCCAGCACTGCGGAGTGGCCAGTGGCAACCTTAC 2460  
DB 2401 TTGCAATTCGCCAGTGCAGTGCACACCCAGCACTGCGGAGTGGCCAGTGGCAACCTTAC 2460  
QY 2461 TTTGCTGCAAGCTCAAGCTGGTGCAGATGCTGTTGACGCTTCCGACCACTGCTGGC 2520  
DB 2461 TTTGCTGCAAGCTCAAGCTGGTGCAGATGCTGTTGACGCTTCCGACCACTGCTGGC 2520  
QY 2521 ACCACCTCCAGCCATCCCTGTGCTGCAATTTGTGCTGCAATTCGCGCACACCCGTCGGAT 2580  
DB 2521 ACCACCTCCAGCCATCCCTGTGCTGCAATTTGTGCTGCAATTCGCGCACACCCGTCGGAT 2580  
QY 2581 ACCGTTTGGAGCTCGAGCTGTTTCTGACCTCGAGCGGTACTGCGGAGCAGTGGCGGA 2640  
DB 2581 ACCGTTTGGAGCTCGAGCTGTTTCTGACCTCGAGCGGTACTGCGGAGCAGTGGCGGA 2640  
QY 2641 CTGTACTCCCACTTTGAGTCTGGAAACCCAGCGCCAAACCGGTGCGGTCTTACCGCACGAA 2700  
DB 2641 CTGTACTCCCACTTTGAGTCTGGAAACCCAGCGCCAAACCGGTGCGGTCTTACCGCACGAA 2700  
QY 2701 ATCCGAGGCGAGTGTGCTGCAACCTGCTGACAGGCGCACCGCACTGGCGCTTGGGAT 2760  
DB 2701 ATCCGAGGCGAGTGTGCTGCAACCTGCTGACAGGCGCACCGCACTGGCGCTTGGGAT 2760  
QY 2761 CGTTTCGAACTCATCGAAGACCACTACGCGCGCTTAAATGAGATGCTGGGACCGCCAAAC 2820

```
Db 2761 CGTTTCGAACTCATCGAAGACATACGACGCGTTAATGAGTGTGGAGCCCAACC 2820
Qy 2821 AAGGTACCCCATCTCTCAAGGTTGTTGGGAGCTCGGACATCCACCTCGTTGTTGGGGT 2880
Db 2821 AAGGTACCCCATCTCTCAAGGTTGTTGGGAGCTCGGACATCCACCTCGTTGTTGGGGT 2880
Qy 2881 GTGATCCAGCAGACTTTGCTGCGGATCCACAAAGTACGACATCCAGACTCTGTCTATC 2940
Db 2881 GTGATCCAGCAGACTTTGCTGCGGATCCACAAAGTACGACATCCAGACTCTGTCTATC 2940
Qy 2941 GGTTCCTGCGCGGAGCTTTGTAACCTCAGGTTGCTGGCCAGAGCCACTGCGCAC 3000
Db 2941 GGTTCCTGCGCGGAGCTTTGTAACCTCAGGTTGCTGGCCAGAGCCACTGCGCAC 3000
Qy 3001 CGCGACTGGAAGGCGCTCGAAGGCAAGGACCTCTGAGGAGTTCCTGAGGAAGAG 3060
Db 3001 CGCGACTGGAAGGCGCTCGAAGGCAAGGACCTCTGAGGAGTTCCTGAGGAAGAG 3060
Qy 3061 CAGGCGCACTCGAGCTGATGATTCCAAAGGAAGTTCGCAATAGCTCAACCGCTCTG 3120
Db 3061 CAGGCGCACTCGAGCTGATGATTCCAAAGGAAGTTCGCAATAGCTCAACCGCTCTG 3120
Qy 3121 TTCCGGAAGCCACCGAAGGTTCTCGAGCACCGTGGCCGCTTCGGCAACACCTCTGCG 3180
Db 3121 TTCCGGAAGCCACCGAAGGTTCTCGAGCACCGTGGCCGCTTCGGCAACACCTCTGCG 3180
Qy 3181 CTGATGATCGTGAATCTTCTACCGCTGCTGCAAGGCGCGAGACTTTGATCCGCTG 3240
Db 3181 CTGATGATCGTGAATCTTCTACCGCTGCTGCAAGGCGCGAGACTTTGATCCGCTG 3240
Qy 3241 CGAGTGTGCGCACCCCACTGTTGTCGCTGATGCGATCTGAGCCAGAGATAAG 3300
Db 3241 CGAGTGTGCGCACCCCACTGTTGTCGCTGATGCGATCTGAGCCAGAGATAAG 3300
Qy 3301 GGTATGCGCAATGTTGTGCGCAACGTCACCGCCAGATCCGCGCAATGCGTGTGCGTAC 3360
Db 3301 GGTATGCGCAATGTTGTGCGCAACGTCACCGCCAGATCCGCGCAATGCGTGTGCGTAC 3360
Qy 3361 CGCTCGTTGAGTGTGTCACCGCAACCGGAGAAAGGAGATTCCTCCAAAGGGCCAT 3420
Db 3361 CGCTCGTTGAGTGTGTCACCGCAACCGGAGAAAGGAGATTCCTCCAAAGGGCCAT 3420
Qy 3421 GTTGTGCAACCATGCTGCTGTTGTCACCGTGAATGTTGCTGAGAGTGAAGTCAAG 3480
Db 3421 GTTGTGCAACCATGCTGCTGTTGTCACCGTGAATGTTGCTGAGAGTGAAGTCAAG 3480
Qy 3481 GCTGAGATGCAAGTCAATCATGAGGCTATGAAAGTGAAGCAACATCACTGCTTCT 3540
Db 3481 GCTGAGATGCAAGTCAATCATGAGGCTATGAAAGTGAAGCAACATCACTGCTTCT 3540
Qy 3541 GTTACCGCAAAATCGATCGCTGTTGTTCTGCTGCAAGAGTGAAGTGGGAC 3600
Db 3541 GTTACCGCAAAATCGATCGCTGTTGTTCTGCTGCAAGAGTGAAGTGGGAC 3600
Qy 3601 TTGATCGTCTGTTTCTTAA 3621
Db 3601 TTGATCGTCTGTTTCTTAA 3621
```

## RESULT 3

```
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
```

```
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

```
Query Match 35.7%; Score 1294.4; DB 3; Length 4403765;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 2099; Conservative 0; Mismatches 1276; Indels 9; Gaps 3;

Qy 238 AAGATCTTGGTAGCAAAACCGCGGCAAAATCGGGTCCGTCCTTCCTCGTGCAGCACTCGAA 297
Db 3317308 AAGGTCTCGTCCCAATCGCGGGAGATCGGATCCGGCTTCCTCGTGCAGCACTCGAA 3317249
Qy 298 ACCGCTGACCCACCGTATTTACCCCGTGAAGATCGGGATCATTCACCGCTCT 357
Db 3317248 CTGGGCTCGGAACCGTGGCCGTTATCGTACGAGGACCGCAATTCGAGCACCGCTCTC 3317189
Qy 358 TTTGCTTCTGAAGCTGTCGCAATGGTACCGAAGCTCACAGTCAAGCGTACCTGGAC 417
Db 3317188 AAGCGGACGATCTTACCAGATCGCGGACATCCGTCACCCGCTGATGATCTGCTG 3317129
Qy 418 ATCGATGAAATTTATCGGTGACGCTAAAGTAAAGACGATGCCATTTACCCGGGATAC 477
Db 3317128 GTGACGAGATCGTGCAGCGCCGCTCGGGGGTGGCGAGCTATCTACCCCTGGCTAC 3317069
Qy 478 GGTCTCTCTGAATAATCCAGCTTSCCGCGAGTGTGGGAAATAGTCTCGGCCATCGGGCG 537
Db 3317068 GGGTTTCTATCGGAGATCCGATCTGGCTGCGGATGCGCGCGCGGCGGATCAGCTTC 3317009
Qy 538 ATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAAACCGCC 597
Db 3317008 GTGCTCCAGCGCCGAGTGTGAGCTGGGTGGGAAATAGTCTCGGCCATCGGGCG 3316949
Qy 598 GCGAAGAGGCTGTGTCAGTTTGGCGGAAATCCACCCGAGGAAACAACTCGATGAG 657
Db 3316948 GCGCGGAAAGCGGCTTGGCCGCTGATGTCCTCGCGCGCTCGGCTCGGTCGACGAA 3316889
Qy 658 ATGTTAAAGCGCTGAAAGCCAGACTTACCCCATCTTTGTGAAGGCACTTGGCGGTGT 717
Db 3316888 CTGCTGTGCTGCGCGCGCATGCCGTTCCGTTGCTGTCAGGCACTTGCCTGCG 3316829
Qy 718 GCGGACGCGGTATGCTTTGTTGCTTCCACTGATGAGCTTCGCAATATAGCAAGAA 777
Db 3316828 GCGGCGCGGATGCTGCTGCTGCGGATATCGCGCGCTTCCGAGGCGATTCGAAGCC 3316769
Qy 778 GCATCTGTAAGCTGAAGCGCTTTCGCGATGCGCGTATATGTCGAACTGCTGTG 837
Db 3316768 GCGACCGGGAAGCCAGTCCGCTTCCGCGGACCCAGCGCTATCTCGAGCAGGCTG 3316709
Qy 838 ATTAACCTCAGCATATTGAAGTGCAGATCCTTTGGCGATCACACTGGAAGTTGTACAC 897
Db 3316708 ATCAATCCAGCCACATCGAGGTGCAATTCGCGGCAAACTCGCGGACGCTGATCCAT 3316649
Qy 898 CTTTATGAAGTCACTGCTCACTGCAAGCTGCTCACCACCAAAAGTTCTCGAAATTCGCCCA 957
Db 3316648 CTTTATGAGGTGACTGAGTGTGCAAGCTGCGCCATCAGAGTCACTGAGCTTGGCGCC 3316589
Qy 958 GCACGATTTGGATCCAGAACTCGGTGATTCGCAATTTGTGCGGATCAGTAAAGTTCTGC 1017
Db 3316588 GCGCGCACCTGAGCGCGGATTCGTTTCAAGATGTCGTCGATGCGGTGCGCTTCGCC 3316529
Qy 1018 CGTCCGATTTGGTACAGGCGCGGNAACGTTGGAATCTTTGCTCGATGAAAGGCAAC 1077
Db 3316528 CGCATATCGGTACAGCTCGCGGCGACCGTCAAGTTCCTGCTGACGAGCGAGGGGAG 3316469
```

QY 1078 CAGCTCTTCATCGAATGAACCCACGATATCCAGGTGAGACACACCGTGACTGAAAGATC 1137  
Db 3316468 TATGTCTTCATCGAATGAATCCCGGGTTAGTGGAGACACACCGTGACCGAGAGATT 3316409  
QY 1138 ACCGAGGTGACCTGTGTGAAGGCGCAGATGCGCTTGGCTGTGGTGTGCAACCTTGAAGGAA 1197  
Db 3316408 ACCGAGGTGACCTGTGTGCGCAGCAGCTGCGCAITTCGCGCGGGGAGACGCTCGAACA 3316349  
QY 1198 TTGGGTCTGACCAAGATAGATCAAGACCCAGGTGACGACTGACAGTCCGCAATCAC 1257  
Db 3316348 TTGGGTCTGCGGACAGAGAGACATCGCACCGATGGCGCGCACTACAGTCCGCGATCAC 3316289  
QY 1258 ACGAAGATCCAAACACGCTTCCGCCACAGATACCGGAACATATCACCGGTACCGCTCA 1317  
Db 3316288 ACCGAGATCCGGCCAAACGCTTCCGCCACAGATACCGGCCGATCAGCGCTTGGCGACC 3316229  
QY 1318 CAGGCGGAGCTGGCTTGGTGTGACGCTGACGCTCAGCTGGTGGCGCAATACCGCA 1377  
Db 3316228 GCGGCGGCTGCCGTGTCCGCTTGAGCGGACGACCAACTGGGCGCAGAAATCACCGCG 3316169  
QY 1378 CACTTTGACTCCATGCTGTGTAATAATGACCTGCCGTGGTTCGACTTTGAAACTGCTGTT 1437  
Db 3316168 TACTTGCATCCATGCTGTGTAATAATGACCTGCCGTGGTTCGACTTTGAAACTGCTGTT 1497  
QY 1438 GCTGTGTCAAGCGCGGTTGGTGTGATTCACCGTGTCTGTGTGTAACCAACATTTGT 1557  
Db 3316108 AGCGGTGCGCGCGGCGGATCGCGGATTCGCGATCCGCGGGGTATCGACGAATATTCG 3316049  
QY 1498 TTCTTGCGTGTGCTGTGCGGAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557  
Db 3316048 TTCTTGCAAGCGGTCTGTGATGATCCCGGACTTCGAGCGGCGCGGTCAACAGTCTTC 3315989  
QY 1558 ATTGCGATCACCGGACTCTCTCAGGCTCAGCTGTGATGATGAGAGGAGCGCATC 1617  
Db 3315988 ATTGATGAGCGGCGGAGCTGCTGACCGCGCGGCTTCGCGGACCGCGGACCAAGATC 3315929  
QY 1618 CTGGAATTAATTGGCAGATGTACCGTGAACAAGCTCATGCTGTGCTGTGCTGCTGCTGCT 1677  
Db 3315928 CTGGAATTAATTGGCAGATGTACCGTGAACAAGCTCATGCTGTGCTGTGCTGCTGCTGCT 1737  
QY 1678 GCAGCTCTATCGATAGCTGCTTAACTCAAGATCTGCGCATGCGACCGGTTCCCGT 1797  
Db 3315868 TACCGGACGCAAGCTGCCGATCTTTGATCTGCGGCGCGACACCGCGCGGTTCAAG 3315809  
QY 1738 GACGCTGAAGAGCTGTGCGGCGGCGGCTTCTGCTGATCTCGCTGAGCAGGACGCA 1797  
Db 3315808 CAGCAGTATGTAAGTTGGGCGCGAAGATTTGCTGTTGGCTGCGGAGTTCGCGCG 3315749  
QY 1798 CTGGAGTTACTGATACACCTTCGCGGATGACACCGATCTTGTGCTGCGACCGGATC 1857  
Db 3315748 CTCGGGCTCACCGATACACCAATTCGCGGATGCTCACCAGTCTGTTACTGCTACCGGAT 3315689  
QY 1858 CGCTATTTCGACCTGAAGCTGTGCGGAGAGGCGCTCGCAAGCTGACTCTGAGCTTTTG 1917  
Db 3315688 CGCAGCAGGACTGTGCGGCTGCGACCGTATCTCGCGGACCATGCGCGAGCTGTG 3315629  
QY 1918 TCCGTGAGGCTGCGGCGGCGGACCTACGATGTGGGATGCGTTCTTCTTGGAGAT 1977  
Db 3315628 TCCGTGAGGCTGCGGCGGCTGCGACTTACGATGTGGCGCTGCGCTTCTCAAGGAGAT 3315569  
QY 1978 CCGTGGGACGCTCGAGCTGCGGAGCGGATGCGGAATGTAAACATTCAGATCTG 2037  
Db 3315568 CCCTGGGAACCGCTGGCCAGCTGCGTGCAGCAATGCGCAATATCTGCTGCGATGCTA 3315509  
QY 2038 CTTCCGCGCGCAACACCGTGGGATACACCGGTACCGAGACTCCGCTGCGCGCGCTTT 2097  
Db 3315508 TTCCGCGCGCGCAATACCGTGGGCTACACCGCGGTACCGGAAATCGTGAAGTCTGCGGCTTT 3315449  
QY 2098 GTTAAGGAGCTGCGAGCTCGCGGCTGACATCTTCGCACTCTTCGACGCGCTTAAAGAC 2157  
Db 3315448 GTGAAGAGCAACAGCAGCGGATTCGATCTTTGATCTTCGACGCGCTCAACAC 3315389  
QY 2158 GTCTCCAGATGCTGCTGAGCAATCGACGATCTCTGGAGACCAACACCGCGGTAGCGAG 2217

Db 3315388 ATCGAGTCAATGCTGCGGCGATTCGACGAGATACCGGAAACAGGTTCTCGATAGCAGAA 3315329  
QY 2218 GTGCTATGCTTTATTTCTGTGATCTCTGTATCCAAATGAAAGCTCTACACCTTGGAT 2277  
Db 3315328 GTGCGATGTGCTACAGGCGGACCTTACCGATCCGGGTGACAGCTGTACACGCTGGAC 3315269  
QY 2278 TACTACATAAGATGCGAGAGAGATGCTCAAGCTGTGGCTGTACATCTTGGCATTAAG 2337  
Db 3315268 TACTACATAAGATGCTGAGCAGATGCTGAGACCGCGGCGCATGCTGCGGATCAAG 3315209  
QY 2338 GATATGCTGCTGCTGCTGCGCCAGCTGCGGTAAACCAAGCTGTGCTACCGCACTTGGCGCGT 2397  
Db 3315208 GATATGCTGCTGCTGCTGCGCCGCGCGCTCAAGGTTGCTCAGCGCTTGGCGAGT 3315149  
QY 2398 GAATTCGATTCGACGATGCTGACACCCAGCACTGCGGCTGCGCACTTGGCGAACC 2457  
Db 3315148 CGCTTCGACCTGCGCTTCACTGTCACACCCAGCAGACACCGGCTGCGCAGCTGCGCAGC 3315089  
QY 2458 TACTTTGCTGCTGCTGCTGCGCCAGCTGCGGTAAACCAAGCTGTGCTGCGCACTTGTCT 2517  
Db 3315088 TATGTCGCGCTTGGCAGCGCGGCGCGATGCGCTGACGCGCGCGCGCTGGCG 3315029  
QY 2518 GGCACCACTTCCAGCCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2577  
Db 3315028 GGAACGACGACGACGCGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3314969  
QY 2578 CATACCGGTTTGAAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2637  
Db 3314968 GACACCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3314909  
QY 2638 GGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2697  
Db 3314908 AAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3314849  
QY 2698 GAATTCGCGGCGGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2757  
Db 3314848 GAGATTCGCGGCGGCGGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3314789  
QY 2758 GATGCTTTCGAACTCATCGAAGACCACTGACGACCGCTTAAATGAGATGCTGCGGACGCGCA 2817  
Db 3314788 GATCGATTCGAAAGATTCGAAAGCGCTTACGCGGCGCGGCGGCTGCTGCGGCGGCTGCT 3314729  
QY 2818 ACCAAGTCAACCCATCTTCAAGGTTGTTGGGACCTTCCGACCTCCACCTGTTGTTGCTGCTGCT 2877  
Db 3314728 GTTAAGTTCACGCGCAACGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCT 3314669  
QY 2878 GGTGTGATCCAGCAGACTTGTGCTGCGGATCCCAAAAGTACGACATCCCAAGACTCTGCTC 2937  
Db 3314668 GGTGTGATCCAGCAGACTTGTGCTGCGGATCCCAAAAGTACGACATCCCAAGACTCTGCTC 3314609  
QY 2938 ATGCGGTTCTGCGGCGGAGCTTGTGTAACCTTCCAGTGTGCTGCGGACGCTGCGGCTGCGGCT 2997  
Db 3314608 CTCGGAATTCGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3314549  
QY 2998 ACCCGGCACTGGAAGGCGCTTCCGAAAGGCAAGCACTCTGAGCGGAGTTCTTCTGAGGAA 3057  
Db 3314548 ACTGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3314492  
QY 3058 GAGCAGGCGCAGCTGCAAGCTGATGTTCCAGGAACTGCTGCAATAGCTCAACCGGCTG 3117  
Db 3314491 GACGAGATTCGCTTATCTGCTGCTGCGGAGCAAG---CGTCAAGGCGGCGGCGGCGGCGGCGGCGG 3314435  
QY 3118 CTGTTCCGAAAGCGCAACCGAAGGTTCTTCCAGGACCGCTGCTGCGGCGGCGGCGGCGGCGGCGG 3177  
Db 3314434 TTATTTCCAGCTGCAAAAGGAAATTCATGAGCAGCGGAGGCTTACGCGGACAGCTGCG 3314375  
QY 3178 GCGCTGAGTATGTTGAATTTCTTACCGGCTGCTGCGGAGGCGGCGGAGACTTTGATCCGCG 3237  
Db 3314374 CAATTTGCGGCGCAACGAGTTCTTCTATGCTGCTGCGGCGGAGGCTGAGAGATCGGCTGAAG 3314315  
QY 3238 CTGCGCAGATGTCGCGACCGCACTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3297



Db	3314314	CTG--GAGCGTGGGGTGGAGCTGTGTGATCGGGCTGGAGGCCAATTTCCGAAACCGGACGAA	3314255
Qy	3298	AAGGGTATGCGCAATGTTGTGGCCAAAGTCACACGGCCAGATCGCCCAATGCGTGTGCGT	3357
Db	3314257	CGGGCATCGAAACGGTGATGTGCATCTCAACGGGCAGCTGGCGGCGGGTCTGTAGTGCGC	3314198
Qy	3358	GACCGTCCGTTGAGTCTGTCCACGGCAACCGCAGAAAAGGCAGATTCTTCCAAACAGGGC	3417
Db	3314197	GACCGAGCAATTGCCAGTCCGTTCCGGCCGCGAGNAGGCCGACCGCGGCATCCCGGA	3314138
Qy	3418	CATGTTGCTGCACGATTTCGCTGGTGTGTTCACCGTGACTGTTCTGCTGAGGTGATGAGTTC	3477
Db	3314137	CACATCCCGCGCCATTTCGCGAGTCTCACCGTGGGGTGTGCGTTCGGCGAGCGGGTC	3314078
Qy	3478	AAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAAACAATCATCTGT	3537
Db	3314077	GGCGCGGCCAACCAACATCGCCACCATCGAGCGCATGAAGATGAAGCCCGATCACCGCC	3314018
Qy	3538	TCGTGTGACGCGCAAAATCGATCGGTTGTGGTTCTGCTCTCAACGAGGTGGAGGTGC	3597
Db	3314017	CCGGTTCGCCGACCCGTGGAGCGGGTGGCGGTGTTCGACACCGGCCAGGTGGAGGGCGGA	3313958
Qy	3598	GACTTGATCGTCGTTTCTCTAA	3621
Db	3313957	GACCTGTTGGTGGTGAGCTGA	3313934

## RESULT 4

```

RESOLI 4
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

	Query Match	35.7%	Score 1294.4;	DB 3;	Length 4411529;
	Best Local Similarity	62.0%;	Pred. No. 0;		
	Matches 2099;	Conservative 0;	Mismatches 1276;	Indels 9;	Gaps 3;
Qy	238	AAGATCTTGGTAGCAACCGCGGGCGAAATCGCGTCCGTGCTTTCCGTGCGAGCACTCGAA	297		
Db	3323034	AAGTGTCTCGTCGCAATCGCGGGGAGATCGGATCCGGGCTTTTCGTGCGCCTACGAA	3322975		
Qy	298	ACCGGTGACGACCGTAGCTATTATCCCGGTGAAGATCGGGATCATATCCACGGCTCT	357		
Db	3322974	CTGGGCGTCGGAACCGTGGCCGTTTATCCGTACGAGGACCCAAATCGAGCACCGTCTC	3322915		
Qy	358	TTTGTCTTGAAGCTGTCCGCAATTGGTACCGAAGGCTCACCAGTCAAGGGGTACTCTGGAC	417		
Db	3322914	AAGCGGACGAGTCTTACCAGATCGGCGACATCGTCAACCGGTGCATACATCTGTGC	3322855		
Qy	418	ATCGATGAAATATTCGGTGCAGCTAATAAAGTTAAAGCAGATGCCATTACCCGGGATAC	477		
Db	3322854	GTCGACGAGATCTGTGCGACGCGCCGTCGGGGGTGCGGAGTATCTACCTTGGGTAC	3322795		
Qy	478	GGCTTCCGTCTGAAAAATGCCAAGCTTCCCGCGGAGTGTGCGGAAAAAGCGATTACTTTT	537		
Db	3322794	GGGTTTCTATCGGAGATCCCGGATCTGGCTCGGCATCGCGGGGGGGGATCAGCTTC	3322735		

1618	CTGGAGTTACTTGGCAGATGTCACCGTGAAACAGAGCTCTATGGTGTGCGTCCAAAGGATGTT	1677
3321654	CTTAACTTCTCGCGCGATGTCAACCGGTCAACAAACCGGTATGGCTCGCGTCCGTTCAACAGATC	3321595
1678	GCAGTCTCTATCGATAAGCTGCTAAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGT	1737
3321594	TACCCGAGCGACAAGCTGCCCGATCTTGATCTGCGGGCGCACCAACCGCGGGTCCAAG	3321535
1738	GACCGCTGGAAGAGCTTGGCCAGCCGGGTTTGCTCGTGATCTCGTGAGCAGAGACGCA	1797
3321534	CAGCAGCTAGTCAAGTTGGGGCCGAAGGATTTGCTCGTTGGCTGCGGAGTCGCGCGG	3321475
1798	CTGGCAGTTTACTGATACCACTTTCGCGCATGCAACCACTTTTCTGTTGCGACCCGAGTC	1857
3321474	GTGCGGGTCAACGATACCACATTCGCGGATGTCCACCACTCGTTACTTGGCTACCCGAGTA	3321415
1858	CGCTCATTTGCGCATGAAGCTCGGCGAGAGGCGGTGCGAAAGTGAACATTCAGATCTTTG	1917
3321414	CGACACAGCGAGCTGTGCGGGTGGCACCGTATCTTCGCGCGACATGCCGACGCTGTG	3321355
1918	TCCGTGGAGCCCTGGGGCGCGCAGCTTACGATGTGGCATGCGTTTCTCTTGAGGAT	1977
3321354	TCCGTGGAGTGTGGGGCGGTGCGACTTACGATGTGGCGCTGCGCTTCTCAAGAGGAT	3321295
1978	CGGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGAACATTCAGATGCTG	2037
3321294	CCCTGGGAACGGCTGGCCACACTGCGTGCAGCATGCCCAATATCTGTTTGAGATGCTA	3321235
2038	CTTTCGCGGCGCAACACCGTGGGATACACCCGTACCCAGACTCGTCTGCCGCGGTTT	2097
3321234	TTGCGGGCGCGAATAACCGTGGGTACTACGCGGTACCGCGAATCGTGACGTGCGGCTT	3321175
2098	GTTAAGGAAGCTGCCAGCTCCGCGGTGGACATCTTCGGATCTTCGACGCGCTTAAACGAC	2157
3321174	GTGCAAGAGCAACAGCCACCGGTATCGACATCTTTGATCTTCGACGCGCTCAACAAC	3321115
2158	GTCTCCAGATCGGTCCAGCAATCGACGAGTCTCTGGAGACCAACACCGCGGTAGCCGAG	2217
3321114	ATCGAGTCGATCGTCCGGCGATCGACGAGTACGGGAAACAGGTTCTCGATAGCAGAA	3321055
2218	GTGGCTATGGCTATTCTGTGTGATCTCTGTATCCAAATGAAAGCTCTACACCCCTGGAT	2277
3321054	GTGCGATGTGCTTACAGGCGACCTTACCGATCCGGGTGAACAGTGTACACGCTGGAC	3320995
2278	TACTACTTAAAGTAGGCAGAGGAGATGTCAGTCTGGCGCTCACATCTTTGGCCATTAG	2337
3320994	TACTACTGAAACTGGTGTGAGCAGATCGTGGACCGCGCGGCCCATGTGCTGGCGATCAAG	3320935
2338	GATATGGCTGCTGTCTTCGCCAGCTGCGGTAAACAAAGTGTGTACCGCACCTGCGCCGT	2397
3320934	GATATGGCTGGAATGTTGCGCCCGCGGCGCTCAAGGTTGTCAGCGCCCTGCGCAGT	3320875
2398	GAATTGATCTGCAGTGCACAGTGCACACCCACAGACACTCGCGGGTGGCAGCTGGCAACC	2457
3320874	CGCTTCGACCTGCCGTTCACTCTGCACACCCACAGACACCGGGTGGCAGCTCGGCAGC	3320815
2458	TACTTTGCTCAGCTCAAGCTGTGAGATGCTGTTGACGGTCTCCGACCACTGTCT	2517
3320814	TATGTGCGCGCTTGGCACCGCGGGCCGATGCCGTGACGCGCGCGCGCGCTGGCG	3320755
2518	GGCACACCTCCAGCCATCCCTGTCTGCGCATTTGTCCTGATTCGGGCACACCGCTGCG	2577
3320754	GGAAACGACAGCCAGCCCGCGCTGAGTCTGATCGTTGCTCCGCTGCCACACCGAGTAC	3320695
2578	GATACCCGTTTGAAGCTTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGC	2637
3320694	GACACCGCCTGTGCGCTTTCGGCGGTGTGCGCCTCTGAGCCGTACTGGGAGGCGTTACGA	3320635
2638	GGACTGTACTGTCATTGAGTCTGGAACCCAGGCCAACCGGTGCGTCTACCGCCAC	2697
3320634	AAAGTGTATGCGCGGTTTCGAGTCTGGGTTGTCGGGGCGCAGCGGGCGGTTTATCACAC	3320575
2698	GAAATCCAGCGCGACAGTTGTCCAACTGCTGTCACAGGCCACCGCACCTGGGCTTGGC	2757

## RESULT. 5

RESOL 3  
US-08-311-731A-1  
; Sequence 1, Application US/08311731A

US-08-311-731A-1  
: Sequence 1. Application US/08311731A

; Patent No. 6583266

; FACILE NO. 0505200  
; GENERAL INFORMATION:

APPLICANT: SMITH. DOUGLAS

APPLICANT: MAQ. TEN-I

APPLICANT: MAO, GEN-I  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

1. TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
2. TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR

TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 411  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
STREET: 600 ATLANTIC AVENUE  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,731A  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: GATES, EDWARD R.  
REGISTRATION NUMBER: 31,616  
REFERENCE/DOCKET NUMBER: C0044/7125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/720-3500  
TELEFAX: 617/720-2441  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32155 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: MYCOBACTERIUM TUBERCULOSIS  
US-08-311-731A-1

Query Watch 34.5%; Score 1249.6; DB 4; Length 32155;  
Best Local Similarity 62.0%; Pred. No. 0;  
Matches 2087; Conservative 0; Mismatches 1274; Indels 13; Gaps 7;  
  
QY 238 AAGATCTTGGTAGCAAAACCGCGGAAATCGGGGTCGGTCTTCCGTGACGACATCGAA 297  
DB 1574 AAGGTCTCTGCGCAATCGCGGGAGATCGGATCCGGGCCCTTTCTGTCGGCCCTACGAA 1633  
QY 298 ACCGTGACGACCGTAGCTATTTACCCCGTGGAGATCGGGGATCTTCCACCGCTCT 357  
DB 1634 CTGGCGCTCGGAACCGTGGCGGTTTATCGGTACGAGGACCGCAATTCGCGACACCGCTCTC 1693  
QY 358 TTTGCTTCTGAAGCTGTCGCGATGTTGATCGGAAGGCTCACCAAGTCAAGCGCTACCTGGAC 417  
DB 1694 AAGCGGACGAGCTTTACAGATCGCGGACATCGGTACCCGGTGCATGATACCTGTGC 1753  
QY 418 ATCGATGAATATTCGGTGACGCTAAAGATTAAGAGAGATGCGCATTTACCGGGATAC 477  
DB 1754 GTGACGAGATCTGTCGCGACCGGCCGTCGGGGGTCGCGACGCTATCTACCCCTGGCTAC 1813  
QY 478 GGTCTTCTGCTCAAAATGCCAGCTTGCCCGGAGTGTGCGGAAACCGGCAATTAATTTT 537  
DB 1814 GGGTTTCTATCGAGATCGGATCTGCTGCGGATGCGCGGCGGCGGATCAGCTTC 1873  
QY 538 ATTGGGCCAACCCAGAGGTTCTTGATCTCAACGGTGATTAAGTCTCGCGCGGTAAACCGCC 597  
DB 1874 GTCGTCCCAAGCGGAGTGTGAGTGTGCTGGGATTAAGTCTCGCGCCATCGCGCG 1933  
QY 598 GCGAAGAGCGTGTGCGGAGTTTGGCGGAATCCACCCGAGCAAAACATCGATGAG 657  
DB 1934 GCCCGGAGCGGCTGCGGCTGTGATGTCCTCGCGCGCTCGGCTCGGTGACGAA 1993  
QY 658 ATCGTTAAACCGCTGAAGCCAGACTTACCCCATCTTTGTGAAGGAGCTTGCCTGGTGT 717  
DB 1994 CTGCTGTCGGTGGCGCGCATGCGGTTTCCGTTTCCGTTTCCGTTCAAGGACGTTGCGCGTGC 2053

QY 718 GCGGACGCGGTATGCGTTTGTGCTTCACTGATGAGCTTGCAGAAATTAGCAACAGAA 777  
DB 2054 GGGGCGCGGGTATGCGTCTGTCGCGATATCGCGGCTTCCGAGCGCATCGAAGCC 2113  
QY 778 GCATCTGTTGAAGCTGAAGCGCTTTCGCGATGCGCGGTATATGTCGAGCGCTCTGTG 837  
DB 2114 GCGAGCGGGAGCGGAGTTCGCGCTTCGCGGACCCGACGGTCTATCTCGAGCGAGTG 2173  
QY 838 ATTAAACCTTCAGCATATTGAAGTCAGATCTTTCGCGATCACTTGGAGAAAGTTGTACAC 897  
DB 2174 ATCATCCAGCGCATCATCGAGTGCAGATTCTGGCGGACAACCTCGCGACGTGATCCAT 2233  
QY 898 CTTTATGAAGCTGACTGCTCACTGAGCGTCTGTCACCAAAAGTTGTGGAATTCGGCCA 957  
DB 2234 CTCTATGAGCGTGAAGTGTGAGCGTTCGCGCTTCGCGGATCATGAGTTCGCGGCC 2293  
QY 958 GCACAGCATTTGATCCAGAACTGGGTGATCCGATTTGTCGGATGTCAGTAAAGTTCTGC 1017  
DB 2294 GCGCGCACTTGAAGCGCGGATTCGTTTACAGATGTCGTCGATGCGGTTCGCTTCGCC 2353  
QY 1018 CGTCTCATTTGGTTACAGCGCGCGGAAACGCTGAGTTCGTTGTCGATGAGAAAGGCAAC 1077  
DB 2354 CGCCATATCGGGTACAGCTGCGCGGACCGCTCGAGTTCCTGCTGAGCAGCGAGGGAG 2413  
QY 1078 CAGCTCTTCATCGAAATGAACCCAGCTATCCAGTTTGAACACACCGTGAAGAAAGTC 1137  
DB 2414 TATGTTTTCATCGAGATGATTCGCGGGTTCAGTTGAGACACACCGTGAAGGAGATT 2473  
QY 1138 ACCGAGTGAACCTGTGAAGGCGAGATGCGCTTGGCTGCTGTCGCTCAACCTTGAAGAA 1197  
DB 2474 ACCGAGTGCACCTGCTCGCGACGCTGCGCATTTGCCCGCGGGAGAGCGCTCGAACAA 2533  
QY 1198 TTGGTCTGACCAAGATGAAGTCAAGACCCAGTTCGAGCTGAGCTGAGTCCGCGCATCACC 1257  
DB 2534 TTGGGCTCGCGGAGGAGGACATCGACCGATGCGCGGCTGAGCTGAGTTCGCGGATCACC 2593  
QY 1258 ACGGAAGATCCAAACACCGCTTCCGCCAGATACCGGAATATATCAACCGGTACCGCTCA 1317  
DB 2594 ACCGAGATCCGCGCAAGCGCTTCCGCCGCGGCTGAGCTGAGTTCGCGGATCAGCGCTT 2652  
QY 1318 CAGCGGAGCTGCGCTTCTGTCGAGTGCAGCTGAGCTGAGTTCGCGGATCAGCGGATCAGCG 1377  
DB 2653 GCGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 2711  
QY 1378 CACTTTGACTTCTGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1437  
DB 2712 TACTTCACTCTCATGCTGCTCAAGCTGACCTGCTGCGGCTGACCTTCCCTACCGCAGTG 2770  
QY 1438 GTCGTGCAAGCGCGGCTTGGCTGAGTTCACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1497  
DB 2771 AGCGTTCGCGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 2830  
QY 1498 TTCTTGGTGTGCTTGTGCGGAGAGGACTTCACTTCCAAAGCGGCTGCGGCTGCGGCTGCGG 1557  
DB 2831 TTCTTGAAGCGGCTTGTGATGACCGCGACTTCCGAGCGGCGGCTGCGGCTGCGGCTTCC 2890  
QY 1558 ATTGCCGATCACCGGCACTTCTGAGTTCAGCTTCACTGCTGATGATGAGGAGGAGGATGCT 1617  
DB 2891 ATTGATGAGCGGCGGCTGCTGACCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 2950  
QY 1618 CTGATTTACTTGGCAGATGTCACCGTGAACAAAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1677  
DB 2951 CTTAACTTCTGCGCGATGTCACCGTCAACAAACCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3010  
QY 1678 GAGCTCTCTATCGAATAGCTGCTTAAATCAAGATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1737  
DB 3011 TACCGGACGACAGCTGCGGCTTCTGATCTGCGGCGCGGCTGCGGCTGCGGCTGCGGCTGCGG 3070  
QY 1738 GACCGCTGAAGCAGCTTGGCGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1797  
DB 3071 CAGCGACTAGTCAAGTTGGGCGCGAAGGATTTGCTGTTGGTGTGCGGAGTTCGCGCGGCGG 3130

QY 1798 CTGGCAGTTACTGATACCACTTCGCGGATGACACACAGCTCTTGTGTTGGACCGGATC 1857  
Db 3131 GTCCGGGTACCCGATACCAATTCGCGGATGCTCAACAGTGTACTGTGCTACCCGAGTA 3190  
QY 1858 CGCTCATTTCCATCTGAAGCCTGCGGCGAGAGCGGTGCGAAAGCTGACTCTCGAGCTTTTG 1917  
Db 3191 CGCACAGCGGAGTGTCTCGCGGTGGCAGCGTATCTCTCGCGGACCATGCGCGAGCTGTG 3250  
QY 1918 TCGTGGAGCCCTGGGCGCGCGGACCTAGCATGTGGCATGCGTTCTCTTTGAGGAT 1977  
Db 3251 TCCGTGGAGTGTGGGCGCGGTGCGATTAAGATGTGGCGTGGCTTTCTCAAGAGGAT 3310  
QY 1978 CCGTGGACAGGCTCGACAGCTGCGCGGCGGATGCGGAAATGTAACATTCAGATGCTG 2037  
Db 3311 CCTTGGGACGCTGGGCCACACTCGGTGCGAGCAATGCCAATATCTGTTGAGATGCTA 3370  
QY 2038 CTTTGGCGCGCAACACCGTGGGATACACCCGCTACCGGATCTCTGCGCGGTTT 2097  
Db 3371 TTGCGGCGCGCAATACCGTGGGTACACCGCTACCGGATTCGCGGATCTGCGGCTT 3430  
QY 2098 GTTAAGGAAGCTGCGAGCTCGGCGGTGGACATCTTCCGATCTTTCGACGCGCTTAAACGAC 2157  
Db 3431 GTGCAAGAGCAACAGCCACCGGTATCGACATCTTTCGATCTTCGACGCGCTCAACAC 3490  
QY 2158 GTCTCCAGATGCTCAGCAATCGAGCGAGTCTCTGGAGACCAACACCGGCTAGCCGAG 2217  
Db 3491 ATCAGTCTGATGCTCGCGGCGATCGACAGTACGCGAAACAGGTTCTCGATAGCAGAA 3550  
QY 2218 GTGCTATGCTTATCTGCTGTGATCTCTCTGATCCAAATGAAAGCTCTACACCTCGAT 2277  
Db 3551 GTCCGATGTCTACACGCGGACCTTACCGATCCGGGTGAACAGCTGTACAGCTGGAC 3610  
QY 2278 TACTACCTAAGATGGCAGAGGAGTCTGCAAGTCTGCGGCTCAGATCTTGGCCATTAAG 2337  
Db 3611 TACTACCTGAACCTGGTGAAGATCTGTGAGCGCGCCCATGTGCTGGCGATCAAG 3670  
QY 2338 GATATGCTGTCTGCTTCGCGGAGTCTGCTGAACCAAGTGTGCTCAACGCTGCGCGCT 2397  
Db 3671 GATATGCTGAAGTGTGCGCGCGCGCGCTCAAGGTTGTGACGCGCTTGGCAGT 3730  
QY 2398 GAATTCGATTCGCAAGTGCAGTGCACACCGACGACACTCGCGGTGGCCAGCTGGCAAC 2457  
Db 3731 CGCTTCGACCTGCGCGTTCACCTGCACACCGACGACACACCGGTGGCCAGCTCGCCAGC 3790  
QY 2458 TACTTTGCTGAGCTCAAGCTGGTGCAGATGCTGTGACGCTGCTTCGCAACCACTGTCT 2517  
Db 3791 TATGTGCGCGTGGCAGCGCGCGCGCGATGCGGTGACGCGCGCGCGCGCTGGCG 3850  
QY 2518 GGCACCACTCCAGCCATCCCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2577  
Db 3851 GGAACGACAGCCAGCCGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3910  
QY 2578 GATACCGGTTGAGCTCGAGCTGCTTCTGACCTGAGCGGCTACTGGGAGCAGTGGC 2637  
Db 3911 GACACCGGCTGCTGCTTTCGCGGTGCTGCGCGCTGAGCGCTACTGGAGGCGTTACGA 3970  
QY 2638 GGAATGCTACCTGCCATTTGAGTCTGGAACCCCGAGCCCGCTGCTGCTGCTGCTGCTGCT 2697  
Db 3971 AAAGTGTATGCGCTTTCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4030  
QY 2698 GAATCCAGGCGGACAGTGTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2757  
Db 4031 GAGATTCGCGGCGGCGCACTGCTCAATCTGCGCGCAGCAAGCAATGCTGCTGCTGCTGCT 4090  
QY 2758 GATCGTTTCGAACCTCATCGAAGACCACTGACGAGCGCTTAAATGAGATGCTGGAGCGCCA 2817  
Db 4091 GATCGATTCGAAGAGATCGAAGAGCGCTTACGCGGCGCGCAGGAGTGTGGGCGAGCTG 4150  
QY 2818 ACCAGGTCAACCATCTCAAGGTGTGGCGACCTCGCATCTCGACCTCGTGTGGCG 2877  
Db 4151 GTTAAGGTCAACCAAGTGTGGCGGATCTGCGGCTGCGACCTGCTGGCGCGC 4210  
QY 2878 GGTGTGGATCCAGCAGACTTTGCTGCGCGATCCACAAAGTACGACATCCAGACTCTGTC 2937

Db 4211 GGTGTGATCGACAGCAATTCGCTCCGATCCAGCGGATTTGGCATCCCGAATCGGTA 4270  
QY 2938 ATCCGCTTCCTGCGCGGAGCTTGTGATCCCTGAGGTGGCTGGCGAGAGCCACTGGC 2997  
Db 4271 CTGGAATTTCTGCGGCGGAGCTGGGTGATCCGCGCGGGTGGCCCGAACCGCTGGC 4330  
QY 2998 ACCCGCCTCTGGAAGCCGCTCCGAGGCAAGGCAACCTCTGACGGAAGTTCTCTGAGGA 3057  
Db 4331 ACTGCGGCTGCGCGGTTCGCGGCGGCGCCAGGCCACT--GCGCAATTGGCGCGGAC 4387  
QY 3058 GAGCAGCGGACCTTCGAGCGTGTGATTCGAGGAAGCTGCAATAGCTCAACCGCTG 3117  
Db 4388 GACGAGATTCCTTATCTGCTGCGTGGAGCCAAAG--CGTCAGGCCACCTCGAAGAGCTG 4444  
QY 3118 CTGTTCCCGAAGCAACCGGAAGTTCCTCGAGCACCGTCCGCTTCGCGCAACACCTCT 3177  
Db 4445 TTATTTCCAGTCCAAAGAGGATTCATGAGCACCGGGAAGCTACGCGGACACGTCG 4504  
QY 3178 CGCTGAGTATGCTGAATTCCTTACGCGCTGGTTCGAGGCGCGGAGACTTTGATCCGC 3237  
Db 4505 CAATTGTGCGCAACACAGTTCCTTATGCTGCGCAAGGTGAAGAGCATCGGCTGAAG 4564  
QY 3238 CTGCGAGATGTCGACACCCCACTGCTTGTTCGCTGATGCTGCTGAGCCAGACGAT 3297  
Db 4565 CTG--GAGCGTGGGTGAGCTGTTGATCGGCTGGAGGCCATTTCCGAACCCGACGAA 4621  
QY 3298 AAGGTATGCGCAATGTTGGCCAAAGTCAACCGGCGAGATCCGCCAATGCTGCTGCT 3357  
Db 4622 CGCGCATGCGCAACCGTGTGTCATCTCAACCGGCGAGCTGCGCGGCTGCTAGTCCGC 4681  
QY 3358 GACCGCTCGTGTGATCTGTCACCGCAACCGCAAGGAGGAGATTCCTCCAAACAGGCG 3417  
Db 4682 GACCGCAGCTTCCAGTTCGCTGCGCGCGCGGAGGAGCGCGGCAATCCCGGA 4741  
QY 3418 CATGTTGCTCACCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3477  
Db 4742 CACATCGCGCGGCTTTCGCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4801  
QY 3478 AAGCTGAGATGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 3537  
Db 4802 GCGCGCGGCGCAACCATCGCCACCATCGAGCGATGAGTGCAGTGCAGTGCAGTGCAGTGC 4861  
QY 3538 TCTGTTGACGCGCAAAATCGATCGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3597  
Db 4862 CGGTTGCGCGCAACCGTGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4920  
QY 3598 GACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3621  
Db 4921 GACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4944

RESULT 6

US-09-134-000C-2302  
; Sequence 2302, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2302  
; LENGTH: 3492  
; TYPE: DNA  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-2302

Query Match	14.1%;	Score 509.2;	DB 4;	Length 3492;
Best Local Similarity	49.1%;	Pred. No. 1.1e-134;	Indels	51;
Matches 1661;	Conservative	0;	Mismatches 1668;	Gaps 10;
QY	233	TCAAAAAGATCTTTGGTAGCAAAACCGCGGGGAAATCGCGTCCGTGCTTTCCGTGACGACAC	292	
DB	65	TGAAAAGGTAATTAGTTGCCAATCGTGGCGAAATTCGGATTGCAATTTTTAGAGCTTGTGA	124	
QY	293	TGGAACCGGTGACGCCACGGTAGCTATTATACCCCGTGAAGATCGGGATCAATCCACC	352	
DB	125	CAGAATTAGATATCCGTACAGTAGCAATTTATGCTGCAAGAAGATGAGTATCTGTGTTCAATC	184	
QY	353	GCTCTTTTGTCTTGAAGCTGTCCGCAATTTGGTACCGAAGGCTCACAGTCAAGCGGTACC	412	
DB	185	GTTTCAAAAGCAGATGAAGCGTATTTAGTTGGTAAAGCGGAAAAAACAATTTGAAGCCTATT	244	
QY	413	TGGACATCGATGAAATTTTCGGTGCAGCTTAAAGAGTTTAAAGCAGATGCCATTTACCCGG	472	
DB	245	TAGACATTTGAAAATATTTATCCAAATTTGCCAAAAAATTCGGGAGCAGATGCCATTCATCTCG	304	
QY	473	GATACGGCTTCTGTCTGAAAAATGCCAGCTTGCCTCCGCGAGTGTGCGGAAAAACGGCAATTA	532	
DB	305	GTATGCTTTTTATCAGAGAACTTAGCGCTTCGCCGAAACGTTGTGAAGAAGAAGGAATCA	364	
QY	533	CTTTTATTTGGCCAAACCCAGAGGTTCTTGATCTCACCGGTGATAGTCTCGCGCGGTAA	592	
DB	365	TTTTTGTGCGACCTTAAACACATCATTTAGATATTTTTTGGCGATAAATTTAAAGCGGAAG	424	
QY	593	CGCCCGCAAGAAGGCTGTCTCCAGTTTTCGCCGAATC---CACCCCGAGCAAAAACA	649	
DB	425	AAGCAGCTGTAGCTGGGGGATTCGTTTCGATTCAGGCTCAGACGGGCCAGTAGCAACGG	484	
QY	650	TCGATCAGATTCGTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTTGTGAAGCGCAGTTG	709	
DB	485	TGGAAGAGGTTGTAGCTTTTGGTGAACACATAGCTTTCCTATCATGATTAAGCTGCTTT	544	
QY	710	CCGGTGTGCGCAAGCGGTATGGTTTTTGTGCTTCACTCATGATGAGTTCGCAAAATTAG	769	
DB	545	TAGTGTGCGCGGTTCGCGGATGCGGCTTCCCAACGATGCCAAGAACGACGAGAAGGTT	604	
QY	770	CAACAGAGCATCTCGTGAAGCTGAAGCGGCTTTTGGCGATGCGCGGTATATGTGCAAC	829	
DB	605	ACGAAAGAGCAAAAGTGAAGCGAAGCAGCCTTTGGTTCTGACGAGGTTTATGTTGAAA	664	
QY	830	GTGCTGTGATTAACCTCAGCATATTGAAGTGCAGATCTTTGGGATCACTATGAGGAAG	889	
DB	665	AGTATATTTCTAATCTCTAAACATATCGAAGTACAAATTTTGGCGCATCATCATGGGAACG	724	
QY	890	TTGTACACCTTTATGAACGTGACTGCTCACTGACAGCGTCTGCACCAAAAAGTTGTGCAAA	949	
DB	725	TCTTTGCAATTTATTTGAACGTGATTTGTTCCGTTCAACGGCGCCACCAAAAAGTGTAGAAG	784	
QY	950	TTGCGCCAGCAGACATTTGGATCCAGAACTCGGTGATCGCATTTGTGGGATGCGAGTAA	1009	
DB	785	TCGACCATATGTATCAATGAATGAAGAAACAGCAGCAGCATTTGTTCGCGCTGCTGTGTC	844	
QY	1010	AGTTCTGCCCTCCATTTGTTTACAGGGCGGGAAACCGTGGAAATCTTGGTTCGATGAAA	1069	
DB	845	AGTTAATGGCACATGTCGGCTACGTGAATGCGGGAAACGGTTGAGTTTTAGT---AGAAG	901	
QY	1070	AGGGCAACAGGCTTTCATCGAAATGAACCAAGTATCCAGTTTGAGTCACACCGTGACTG	1129	
DB	902	GGGATCAGTTTTATTTATTTGAAGTGAATCCTCGTGTTCAGGTAGAACATATCATATCAACG	961	
QY	1130	AAGAAGTCAACGAGGTGGACCTGGTGAAGCGCAGATGCGCTTGGCTGTGGTGCACACCT	1189	
DB	962	AAATGATTAACAGATATTGATATTGTGATCTCTCAATTAACAATTTGCGCAAGGGCTTGATT	1021	
QY	1190	TGAAGGAA-----TTGGGTCTGACCCAGATATAGATCAAGACCCACGGTGCAGCACTGC	1243	
DB	1022	TGCATTAAGATATGCAATTTGCCAAAAACAAACGAATTTGACATTTAAAGGCGCGCTATTTC	1081	
QY	1244	AGTGCCGCATCACCAACGGAAGATCCAAACAAACAGCGCTTCCGCCCGAGATACCGGAATATCA	1303	

DB	1082	AATGTCCGATTACGACTGAAGATCCCCTGAACAGTTTATGCCAGACACAGGGAATCG	1141
QY	1304	CCCGTACCGCTCACAGGGGAGCTGGGTTCTCT---TGACGCTGAGCTCACTCG	1360
DB	1142	ATACGTATCGTTCCACAGGTGTTTTGGTGGGTTAGATGATGAGGAATGCTTATCTG	1201
QY	1361	GTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAATGACCTCGCGTGGTTCG	1420
DB	1202	GCTACGCTGACTCCCTACTTTGATCTTTATTTGGTTAAAGTCTGACGATGGTTCT	1261
QY	1421	ACTTTGAACTGCTGTGCTGCTGCAACGGCGGTTGGCTGAGTTCACGGTCTCTGGTG	1480
DB	1262	CTTTTGAACAGGCAATCAGTAAATCAACGCTGCTTAAAGAAATTTGGATTCGTTGGG	1321
QY	1481	TTGCAACCAACATTTGTTCTTGGCTGGTGGCGGAGAGGACTTCACCTTCCAAAC	1540
DB	1322	TGAACAACAATATTCGTTTTTACAAATTTCTGAGCTATCCAGCGTTTCAATCTGGG	1381
QY	1541	GCATGCGCACCGGATTCATTCGCGATCACCGCACCTCTCTTCAAGGCTCACCTGCTGATG	1600
DB	1382	AAGCCAAAACAACCTTTATGATAATACACCTGAATTTATTTGAATTTCTCTGATCGCG	1441
QY	1601	ATGAGCAGGAGCAGATCTCGATTACTTGGCAGATGTCAACGTTGAACAAGCCCTCATGGTG	1660
DB	1442	ATCGTGCAATAAACAATGAATATCATTTGAGAGATGAAGTCAATGGTTCCTCTGGCA	1501
QY	1661	TGCGTCCAAAGGATTTGCACTCTTATCGATAA-----GTCCTTAACATCAAGATC	1714
DB	1502	TTGAAGAAGCGAGAAAAATATTTTGAAGCCCGCGCGTCCGACAGATATTTGAAGTTT	1561
QY	1715	TGCCATCTCCACCGGTTCCGTTGACCGCTGAAGCAGCTTGGCCCGCGCTTGTCTC	1774
DB	1562	CAGAAAAGTGTATCAGCTAAAAATATTTCTAGACGCTCAAGAGCAACCGCGCTCATCG	1621
QY	1775	GTGATCTCCGTGAGCAGCAGCTGCGATTAATGATACCACTTCGCGATGCAACAC	1834
DB	1622	ATTGGGTTAAGAAATCAAGAAAGTGTATGATGACAGTACACAGCTTCGCTGATGCGCAC	1681
QY	1835	AGCTTTGCTTGGACCGGAGTCGCTCATTTCCACATGAAGCTTGGCGGAGAGCGCGTGC	1894
DB	1682	AAAGTTATTAGCTACTCGTGTGAGAACACAGATTTTAAAGCAATTTGCTGCTTAACTG	1741
QY	1895	CAAAAGTGAATCTCTGAGCTTTTGTTCGTGAGCGCTGGGGCGGCGGACCTACGATGTGG	1954
DB	1742	ATGCAGCCCTGCTGAGCTGTTTCTAGTGAATTTGGGGTGGCGCTACGTTCTGATGTTG	1801
QY	1955	CGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTCGCGAGGCGATGC	2014
DB	1802	CCTATCGCTTCTTAAACCGAAGACCCATGGCAACGTTTAAAGAAAATTCGCTGATGTC	1861
QY	2015	CGAATGAAACATTCAGATCTGCTTCGCGCGCAACACCGTGGGATACACCCCGTACC	2074
DB	1862	CAACACACTTTTACAAATGCTGTTCAAGGATCTAATGCGTAGGCTATCAAAATATC	1921
QY	2075	CAGACTCCGTGCGCGCTTTGTTAAGGAGTGCAGCTCCGCGGTGGACATCTTCC	2134
DB	1922	CTGATAATGTTTATTAAGAAATTTTAAAGAAATCCGCTCGCCAAAGGAGTGCATGTTTCC	1981
QY	2135	GCATCTTCGAGCGCTTAAACGAGCTCTCCAGATGCGTCCAGCAATCGACGAGTCTCTGG	2194
DB	1982	GAATTTTGTAGTGTAACTGGATTTCTCAATGGAAGAAAGTATTTCAAGTCTGTCGGG	2041
QY	2195	AGACCAACACCGCGGTAGCGAGTGGCTATGCGTTATTTCTGTTGATCTCTCTGATCAA	2254
DB	2042	ATACCGGAA---AAATTTGCGAAGCAGCAATTTGTTATCTGCGGACATCAATGATCCAG	2098
QY	2255	ATGAAAAGCTCTACACCTGGATTACTTACCTAAGATGGCAGAGGAGATCGTCAAGTCTG	2314
DB	2099	CCCGAGCAAAATATATGTTTCAATTTACCTTGTATGGCTAAAGATTTGGAATTTAG	2158
QY	2315	CGCTCTCATCTTGGCCATTAAAGGATGCTGCTGCTTTCGCCCGAGCTGCGGTAAACCA	2374

Db 2159 GTGCAAAATCATTTGGATTAAGATATGGCTGGCTTATTGTAAGAACCAACAGCTGCTTATC 2218  
Qy 2375 AGCTGTGTACCGCACTGCGCGCTGGAATTCGATCTGCGAGTCAGCTGCACACCCACGACA 2434  
Db 2219 GTTTAATTAAGTAATTAAGGACGACCGAATTTACCAATTCATCTCCACACTCACGACA 2278  
Qy 2435 CTGCGGTGCGCAGCTGGCAACCTACTTTGCTGCGAGCTCAAGCTGGTGCGAGATGCTGTG 2494  
Db 2279 CTAGTGGCAATGGATGATCATACATATCTAGCAGCTACTAAAGCGGCGGTGTATATTTG 2338  
Qy 2495 ACGGTGCTTCGCGACCACTGTCTGGCACCACCTCCCGAGCCATCCCTGTCTGCGCATTTG 2554  
Db 2339 ACGTCGAATGAGTCGATGAGTGGCGCACTAGCCAGCCAAAGTATGAACAGCTCTATATT 2398  
Qy 2555 CTGATTCGCGCACACCGCTGCGGATACCGGTTTGAGCTCGAGCTGTTTCTGACCTCG 2614  
Db 2399 ATGCTTTAGTCAATGTTGAAGCGACGCCAACTATTAACTATTGATATGCACAGAAATCA 2458  
Qy 2615 AGCGTACTGGGAAGCAGTGGCGGACTGTACCTGCTGCTGCTGGAACCCAGGCC 2674  
Db 2459 ATCATTTATGGGAAGATGGCGCATGATTTATCAACATTTGAAATGGTTAAATGCC 2518  
Qy 2675 CAACCGTCCGCTCTACCGCACGAATCCAGCGGACAGTTGTCACACCTGCGTGAC 2734  
Db 2519 CGCAACAGAGTCTATATGATCATGAATGCTGTGGGCACTACTCTAATCTACAGCAGC 2578  
Qy 2735 AGCCACCGCACTGGGCTTTGCGGATCGTTTCGAACCTCATCGAAGCAAACTACGACGCG 2794  
Db 2579 AAGCAAAAGCGGTAGGCTTAGGCGACCGTTGGATGAATCAAAATCAAAATATATCACACAG 2638  
Qy 2795 TTAATGAGATGCTGGAGCGCCAAACAGGTCAACCCATCCCTCAAGTTGTTGGGACC 2854  
Db 2639 TGAATTTGATGTTTGGCGATATTTGTAAGAGTACGCCATCTTCAAAAGTGGTGGGAGACA 2698  
Qy 2855 TCGCACTCCACCTGCTTGGTGGCGGTGTGATCCAGCAGACTTTGCTGCGATCCACAAA 2914  
Db 2699 TGGCACTCTTCATGTTCAAAATTAATCTGACAGACAAGATGTTTATGCGCGTGGTGAAG 2758  
Qy 2915 AGTACGACATCCAGACTCTGTGATCGCTTCTGCGCGGAGCTGTGTAACCTCCAG 2974  
Db 2759 AACTAAGTTTCTGTAATCATGTTGTGACTTTTTCGAAGTGTATTTAGGTCAACCACTG 2818  
Qy 2975 GTGCTGGCAGAGCCACTGGCACCGCGCATCTGGAAGCGCGCTCGAAGGCAAGGCAC 3034  
Db 2819 GTGGTTCCCGAAGAGCTGCAAGATTTCTTAAAGTTCGACCAAGCTTACAGAGC 2878  
Qy 3035 CTCTGACGGA---AGTTCCTGAGAAAGAGCAGGCGCACCTCGAGCTGTATGATTCAGG 3091  
Db 2879 GGCCAGGTGACTTAGCAGCACTGTTGATTTTGGAAAGTACAGAGATTTAGCCGAA 2938  
Qy 3092 AACCTCGCAATAGCTCAACCGCC-----TGCTGTTCCCGAAGCCAA 3133  
Db 2939 AATTTGGTATCAACCTTAATTTGGAAGAAGTTTAAGTTATTAAATGATCCACAGTAT 2998  
Qy 3134 CCGAAGAGTTCTCTGAGCAGCGCTCGCGCTTCGGCAACACCTTCGCGCTGATGATGGTG 3193  
Db 2999 TTTTAGAGTATCGTCAAAAATACGAAACCTTTGGAGATATAACCTTATTAGATACGCCAA 3058  
Qy 3194 AATTTCTTACGGCTGTGCAAGCGCGGAGACTTTGATCCGCTCGCAGATGTGGCA 3253  
Db 3059 CATTTCTTAATGGTATTCGTCAAGA---GAGACATTTGGAAGTTCAAAATTTGAAGAGGAA 3115  
Qy 3254 CCCCACCTGTTTTCGCTGGATGCGATCTCTGAGCCAGACGATPAAGGGTATGCGCAATG 3313  
Db 3116 AAACGTTAATTTATTCGTTTAGATGAATTTGGCGAACTCTGATATTGACGGAACCGTGTCT 3175  
Qy 3314 TTGTGGCAACGTCACCGCCAGATCCGCCCAATGCGGTGGTGGTGGCGCTCCGTTGAGT 3373  
Db 3176 TGTCTTCAATTTGAATGGGCAACGTCGTGAGGTTTATGTTAAAGATGCTCTATTAAAGT 3235  
Qy 3374 CTGTACCGCAACCGCAGAAAGGAGATTCCTCCAAAGGGGCACTGTTGCTGCAACCAT 3433  
Db 3236 CTGCTGTTTCAAGTAAAAACAAAGGCGGAACCAACTAATAAAGAACAAATCGGCGCAACGA 3295

Qy 3434 TCCTCG---TGTTTCACCGTCACTGTTGCTGAAGTGTGATGAGTCAAGGCTGGAGATG 3490  
Db 3296 TGCTGTTCTGTTCTTACAGTATTGCTCAAGCTGGCATAAAGTAGAAAGAGTCAAGC 3355  
Qy 3491 CAGTCGCAATCATCGAGGCTTATGAAGATGAAGCAACAATCACTGCTTCTGTTGACGGA 3550  
Db 3356 CGTTGCTGATTACAGAAGCCATGANAATGGAACGACTATCGAAGCACGTTTTGCGGGA 3415  
Qy 3551 AAATCGATCCGTTGTGGTT 3570  
Db 3416 CGGTAGATCATTTATGTT 3435

RESULT 7  
US-09-134-001C-591  
; Sequence 591, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 591  
; LENGTH: 3465  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-591

Query Match 13.4%; Score 486.2; DB 4; Length 3465;  
Best Local Similarity 48.4%; Pred. No. 3.9e-128;  
Matches 1661; Conservative 0; Mismatches 1718; Indels 54; Gaps 9;

Qy 231 ATTCAAAAGATCTTTGGTAGCAACCCGCGGGAATCGCGTCCGTCGTTCCGTGACG 290  
Db 24 AATAAGAAATTAATCTGTTGCTAACCGTGGTGAATCGCCATTAGAAATTTTAGAGCGGC 83  
Qy 291 ACTGAAACCGGTGCGACGCTAGCTATTATCCCGTGAAGATCGGGATCATTTCCA 350  
Db 84 AGCAGATTAATATCATGACAGTAGCAATTTATCTAATGAAGATAAAGTTCGTTACA 143  
Qy 351 CGGCTCTTTTCTGAGCTGTCGCAATTTGGTACCGAAGGCTCACGAGTCAGGCGTA 410  
Db 144 TAGATATAAGCAGATGAATCCTATCTAGTTGGAAGTGAATTTAGGACCTGCTGAAAGTTA 203  
Qy 411 COTGACATCGATGAATTAATTCGGTGCAGCTTAAAGTAAAGCAGATGCCATTACCC 470  
Db 204 TTTGAATATGAAAGTATCATCGAAGTAGCTTCGCGCAGGTGTCGATGCAATTCATCC 263  
Qy 471 GGGATACCGCTTCTGCTGTAATAATGCCCAGTTGCGCGGAGTGTGCGGAAACGGCAT 530  
Db 264 TGGGTATGGTTTTTAAAGTGAATAAAGTAAACATTTGACGCGGATGTCGTGAGGAGCAT 323  
Qy 531 TACTTTTATGCGCCAAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGGCT 590  
Db 324 TAAATTTATAGTCCGCACTTTGAACATCTAGACATGATTTGGAGATTAAGTTAGGCTAG 383  
Qy 591 AACCCCGCGAAGAGGCTGTGCTGCCAGTTTGGCGG---AATCCACCCCGAGCAAAA 647  
Db 384 AACAATGCTATTAAACGCTAACTTACCTGTAAATCCCGGTTACAGATGCTCTCTTTGAAG 443  
Qy 648 CATCGATGAGTCGTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGAGT 707  
Db 444 TTTGAAGTCGAGAACAGTTTGTCTTAATGAAGAGGTTTACCCACTTATGATTAGGCCAC 503  
Qy 708 TGCCCGTGTGGCGGACGCGGTATGCTGTTTGTGTTCTTCACTGATGAGCTTCGCAAAAT 767

Db 504 AAGCGTGGCGTGTAAAGGTATCGAATCTCGTGAATCAAGCGAATTAGAAGCGC 563  
Qy 768 AGCAACAGAAAGCATCTCGTGAAGCTGAAGCGCTTTCCGCGATGCGCGGTATATGTCGA 827  
Db 564 TTTCCATCGTGGCAATCAGAAAGCGAAAGTCATTTGGTAATAGCGAAGTTTATATCGA 623  
Qy 828 AGTGTCTGTGATTAACCCCTCAGCATATTGAAGTGCAGATCTTTGGCGATCACTGGAGA 887  
Db 624 AAGATATATTGATAATCAAGCATATAGAGGTTCAAGTTATTGGTGTGAATTCGGGAA 683  
Qy 888 AGTTGTACACCTTTATGAAGCTGACTGCTCACTGCGAGCTGCTCAACCAAAAAGTTGTGCA 947  
Db 684 TATCATCTATTGTTATGAAGAGATGCTTCCGTCAACAGCTCATCAAAAGTTTGTGGA 743  
Qy 948 AATTGGCCGACGACAGCATTTGGATCCAGAACTGCGTGTATCGCATTTGTGGCGATGCACT 1007  
Db 744 AGTTGCACCTTCAGTAGGCTTTCTTAAACAAATTAAGAGAGCGAAATTTGTGTGCGCGCAAT 803  
Qy 1008 AAGTTCTCGCGCTCCATTTGGTTTACAGGGCGGCGGACCGTGAATCTTTGGTTCGATGA 1067  
Db 804 TCACCTGATGGAAATATAAATATAGCTCAACGCTGGAACAGTAGAATTTTGTAGTTTCG- 862  
Qy 1068 AAGGGCAACACCGTCTTCATCGAAATGAACCAAGTATCCAGAGTTGAGCAACACCGTGAC 1127  
Db 863 --GGGATGAATTTTCTTCAATGAGGTTAATCCAGTGTTCAGTTGAGCATACAAATTAC 920  
Qy 1128 TGAAGAGTCAACGAGGTGCACTGTTGAAGCGCGAGTATGCGCTTTGGCTGCTGGTGCAC 1187  
Db 921 TGAATATGATTAAGTGTATAGCATTTGTGAAGACGCAAAATTTTGTGCTAATGGAGAATC 980  
Qy 1188 CTTGAAGGAATTTGGGCTGACG-----CAAGATAGATCAAGACCAACCGGTGCAGC 1238  
Db 981 GTTATTTGGAGATAAATCTCTATGCCACAGCAAAATGAATTCAAACATTAAGGTATGC 1040  
Qy 1239 ACTGAGTGGCGATCACACGGAAGATCCAAACAAACGCGCTTCGCGCCAGATACCGGAAC 1298  
Db 1041 GATACAATGTCTGTATAAACAACCTGAAGATCTCTAATGATTTTATGCCAGATTTCTGGCAC 1100  
Qy 1299 TATCACCGGTACCGCTTCAACGAGGCGAGCTGGCGTTGCTGTTGAGCGTGCAGCTCAGCT 1358  
Db 1101 AATTATGATATCATGATCAAGTGGCGGTTTGGTGTGAGACTTGATGCGGGATGGAAT 1160  
Qy 1359 C---GGTGGCAATCACCGCACATTTGACTCCATGCTGTGGTGAATAAGACTGCGCGTGG 1415  
Db 1161 CCAAGGTGCAGAAATTTCACTTACTACGATTTCACTATTAGTTAAGCTTTCTACACATGC 1220  
Qy 1416 TTCGGACTTTGAACCTGCTGTGCTGTCGACAGCGCGGTTGGCTGAGTTCAACCGTGC 1475  
Db 1221 CGTTTCATTTAAACAGCTGAAGAGAAATGGAACTTCATTACCGGAAATGCGGAATTCG 1280  
Qy 1476 TGGTGTGTCAACCAACATTTGGTTCTTGGTGGTGTGCTGGGGAAGAGACTTCACATTC 1535  
Db 1281 TGGCGTAAAGACGAATATTCCATTTCTCATCAATGTTATGCGTAATGATATAATTTAGAAG 1340  
Qy 1536 CAAGCGCATCGCACCGGATCTATGCCATCACCGCACTCTTCAAGCTCCACCTGC 1595  
Db 1341 TGGTGAATTAATAAATTTATTGAAGAAACACCTGAACCTTTTCGATTTGCAACCGAC 1400  
Qy 1596 TGATGATGACGAGGACGCACTCTGATTAATTGTCAGATGTCAACCGTGAACCAAGCTCA 1655  
Db 1401 ATTGGACAGAGGTACCAAGACTTTAGAGTATATTGTAATGTGACGATAAAGGATTTCC 1460  
Qy 1656 TGGTGT-----GCGTCCAAAGATGTTGCGAGTCTCTATCGATAGCTGCGCTTAACATC-- 1707  
Db 1461 TAATGTAGAAAAGCGTCCAAAACCAAGATATGAATCTACCAAAATTTCCAAAATTTCTCA 1520  
Qy 1708 -AAGGATCTGCCACTGCCAGGGTTCCCGTGACCGCTGGAAGAGAGCTTTGGCCACGCGC 1766  
Db 1521 AAGAAATCAATCACTATTATTGGACACAAACAAATTTCTTGAGCAACATGGACCAACAG 1580  
Qy 1767 GTTTGCTGCTGATCTCCGTGAGCAGACGCACTGGCAGTTACTGATACACCTTCCGCGA 1826

Db 1581 TGTTACAAATTTGGTTAGAGAACAGAGATGTTTTAATTACCGATPACTACATTTAGAGA 1640  
Qy 1827 TGCAACCACTGCTTTGCTTGGACCCGAGTCCGCTCATTTCCGACTGAAGCTTCGCGCAGA 1886  
Db 1641 TGCAACCACTGCTTTGCTTGGACACAGTGTAGAGAACAAAGATATGATGAACATTGCATC 1700  
Qy 1887 GGGCGTGGCAAGCTGACTGCTGAGCTTTTGTCCGTGGAGGCTTGGGGCGCGCGACCTTA 1946  
Db 1701 TAAAACTGCTGGAAGTTTTTAAAGATAGTTTTTCAATTAGAAATGTGGGGTGGTGCACATT 1760  
Qy 1947 CGATGTGGCGATCGCTTTCTCTTTGAGGATCCGTCGGGACAGGCTCGACGAGCTGGCGA 2006  
Db 1761 TGAATGCGCTATAATTTCTTGAAGAGAAATCATGCGGAACGTTTGAAGATTTGCGCA 1820  
Qy 2007 GGGGATGCGGAATGTAAACATTCAGATGCTGTTCCGCGCGCAACACCGTGGGATACAC 2066  
Db 1821 AGCCATTCGGAATGTGTTATTCGAATGTTATTACGAGCTTCGAACGCACTAGTTTATAA 1880  
Qy 2067 CCGGTACCCAGACTCGCTCTCGCGCGTGTGTTAAGGAAGCTGCCAGCTCCGGCTGGA 2126  
Db 1881 AAACATCTCTGATTAATTAAGAAATTCGTTATGAAGTGCAAAAGCTGGTGTAGA 1940  
Qy 2127 CATCTTCGCACTCTCGACGCGCTTAACGAGCTCTCCAGATGCGTCCAGCAATCCACGC 2186  
Db 1941 TGTTTTCCGTATATTCGACTCATTTGAATGAGTAAATGAAAGTAGCGAATGAAGC 2000  
Qy 2187 AGTCTTGGAGACCAACACCGCGGTAGCGAGTGGCTATGCGTTATTTCTGTTGATCTCTC 2246  
Db 2001 TGTTCAAGAGCTGGAATGTTATCTGAGGTACAATTTGCTATACAGGTGATATTTTAA 2060  
Qy 2247 TGATCCAAATGAAGAGCTCTACACCCCTGGATTAATCTAAAGATGGCAGAGAGATCGT 2306  
Db 2061 TGCTGAAGCTTCCAATATTTATATTTAGATTATACGTTTAAATGGCTAAAGACTGGA 2120  
Qy 2307 CAGCTTGGCGCTCACTCTTGGCCATTAAGATATGGCTGGTCTGCTTCGCCCGCTGC 2366  
Db 2121 AAGGAAGGATTCATATATTTAGCAATTAAGATATGGCTGGTGTATTGAAACCGAAGC 2180  
Qy 2367 GGTAAACCAAGCTGGTCAACCGACTGGCGCGTGAATTCGATCTGCCAGTGCACGTCACAC 2426  
Db 2181 AGCTTAGCAATTAATTTGGTGAATTTAGTGAAGCAACACATCTTCCAATTCATTTACATAC 2240  
Qy 2427 CCAGCACTTGGCGGTGGCCAGCTGGCAACCTACTTTGCTGAGCTCAAGCTGGTGCAGA 2486  
Db 2241 ACATGATACTAGTGGAAATGGATTTGTGACATATAAACAGCAATTTGATCTGGCGTAGA 2300  
Qy 2487 TGCTGTTGACGGTGTCTCCGACCACTGTCTGCAACCACTCCCGAGCCATCCCTGCTGC 2546  
Db 2301 TATTATAGATCTGCTGTTGCTATGATGTTTAAAGAGTCAACCAAGTGCATAATTC 2360  
Qy 2547 CATTTGTTGCAATTCGCGCACACCCGTGCGGATACCGGTTTGAGGCTTCGAGGCTGTTTC 2606  
Db 2361 ATTATATTATGCACTAAATGGAATTTCCACGTAATTTAAGAACTGATATTGATGGGTTAGA 2420  
Qy 2607 TGACCTCGACCGCTACTGGGAAGCAGTGGCGGACTGTACTCGCCATTTGAGTCTGGAAC 2666  
Db 2421 AGAGTTGAGTCATTTGCTGCTGTAGTCAGACTTACTATGAGACTTTTGAGAGTGTAT 2480  
Qy 2667 CCCAGGCCAACCGGTGGCGTCTACGCCACCAAGAAATCCCAAGCGGACAGTTGTCACCT 2726  
Db 2481 CAAATCACCAAATACAGAAATTTTCAACATGAATGCGCAGGTGGCCAAATTTCAAACCT 2540  
Qy 2727 GCGTGCAGCGCCACCGCACTGGCGCTTCGGATCGTTTCGACTCATCGAAGACACTA 2786  
Db 2541 AAGTCAACAGCTAAAGTTTAGGATTTGGCGCAACGTTTTGATGAAGTCAAGAGATGTA 2600  
Qy 2787 CGCAGCGGTTAATGAGATGCTGGGACGCCCAACCAAGGTCAACCCATCTCTCAAGGTTGT 2846  
Db 2601 TCGTGTGTCAACTCTCTGTTTGGAGATCTTGTAAAGTAACCACTCTTCAAAGGTAGT 2660  
Qy 2847 TGGCGACTCGCACTCCACCTCGTTGGTGGGTGGATCCAGCAGACTTTCTCTGCCGA 2906  
Db 2661 TGGAGATATGCACTATATATGTTGCAAAATGATCTTGATGAAGATACGCTCATCAATGA 2720



2907 TCCACAAAGTAGACATCCAGACTCTGTATCGCGTTCCTGCGCGGAGCTGTGTAA 2966  
Db TGGTTATTAATTTAGATTTCCAGAAATCTGTGTGTCTTTAAAGGTGACATTTGGACA 2780  
Qy CCTCCAGTGTGCTGGCAGAGCCACTGGCAGCCGCGCTGTGAAGCGCGCTCCGAGG 3026  
Db ACCTGTCAACGGATTTCAACAGAAATTTGAGATGTTTATTTAAAGGACAGCAACAT 2840  
Qy CAAGGCACCTCTGACGGAAGTTCTTGAGGAAGAGCAGCGCACTCGACGCTGATGATTC 3086  
Db TACTGAAAGACACAGTGAATACTTGAGCGCGGTGATTTGAAGCAATCCGTCAAGATT 2900  
Qy CAAGG-----ACGTGCAATAGCCTCAACCGCTGCTGTTCC 3125  
Db AAGCGACATACAAAGACGAGTTACAGAACAGATTAATAGTTATGTTACTTTATCC 2960  
Qy GAAGCCAAACGAAGAGTTCTCGAGCACCGTCCGCTTCGCAACACCTCTGCGCTGA 3185  
Db GAAGGTATATAACAATAATTCAAGCAAGAGCAATTTGGTAATGTATCTTTACTGGA 3020  
Qy TGAATGTGAATTTCTACGCGCTGTGCAAGCGCGAGACTTTGATCCCGCTGCCAGA 3245  
Db TACACGCAATTTCTATTTGGCATGCTAATG---GTGAACAGTTGAAATTTGA 3077  
Qy TGTGCGCACCCCACTGCTTGTTCGCTGATGCGATCTCTGAGCCAGACGATAGGGTAT 3305  
Db TACTGTAAACGCTCTAATTTAATTAATTAAGAAATCACTGCAACAGATGAAATGTAA 3137  
Qy GGCATATTTATTCAGCTATGAATGTTCAAGACGATCTATATTTCAAGATGAAA 3197  
Qy CGTTGAGTCTGTCACCGCAACCGCAGAAAGGAGAGTCTCTCAACAGCGCATGTTGC 3425  
Db TGTAAACGAATCTAATGTTAAACCTAAGCGGATTAATCAATCCAATCATATTGG 3257  
Qy TGACCATTCGCTGGTGTGTAC---CGTGACTGTTGCTGAGGTGATGAGCTCAAGC 3482  
Db TGCTCAATTCCTGCTGTCTAAGTCAAGTCAAGTGTCTGAGGGGATGAAGTTCAAGC 3317  
Qy TGGAGATGAGTGCAGTATCAAGGCTATGAAGATGGAAGCAACATCACTGCTTCTGT 3542  
Db TAATCAGCAATTAATTAATCACTGAGCAATGAAGATGGAACGAGATTTCAGGCACCAT 3377  
Qy TGACGCAAAATCGATCGCTGTGTGTTCTGCTGCAACGAGGTGGAAGGTGGGACTT 3602  
Db TGATGGAATTTAATAACAAATCAATGTTGCTAATGAGATGCCATTTGCCACAGGATTT 3437  
Qy GATCGTCTGCTT 3615  
Db ATTAGTGAATTT 3450

RESULT 8  
US-08-956-171E-324/c  
; Sequence 324, Application US/08956171E  
; Patent No. 6593114  
; GENERAL INFORMATION:

APPLICANT: Charles Kunsch  
Gil H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5256  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA

ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171E  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 324:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5030 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 324:  
US-08-956-171E-324

Query Match 13.3%; Score 479.8; DB 4; Length 5030;  
Best Local Similarity 48.3%; Pred. No. 3.3e-126;  
Matches 1657; Conservative 0; Mismatches 1722; Indels 54; Gaps 9;  
Qy 231 ATTCAAAAGATCTGTAGCAAAACCGCGCGGAAATCGCGGTCGTTCCGTGCGAGC 290  
Db AATAAAAAGTTACTTGTCTAACCGGTGAGAAATTCGAAATTCGATATTCAGAGCGC 3787  
Qy 291 ACTGAAACCGGTGCACCGCAGCTAGTATTATCCCGGTGAAGATCGGGATCAATCCA 350  
Db GGCAAGATTAGACATCAGCAGAGTTGCAATTTATCGATGAAGACAAAGATTCATACA 3727  
Qy 351 CCGCTCTTTTCTTGAAGCTGTCCGATGTGACGAAAGGCTCACCAGTCAAGGGGTA 410  
Db TAGATATAAAGCAGATGAATCCTATTATTAGTTGAAGTGATTTAGGTCCTCTGAAAGTTA 3667  
Qy 411 CCTGCATCGATGAATATTCGTGCGAGCTTAAAGATTTAAAGCAGATGCCATTTACCC 470  
Db TTTAAATATTGAGCGTATCATTTAGTAGCAAAACAGCGAATGTGATGCGATTATCC 3607  
Qy 471 GGGATACGGCTTCTGTCTGAAATGCCAGCTTGCCTGCGAGTGTGCGGAAACCGCAT 530  
Db TGGCTATGGAATTTTAAAGTGAATGAACAATTTGCGCGTCTGTCAGAGAGGAAT 3547  
Qy 531 TACTTTATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGGGT 590  
Db TAAATTTATTGGTCTCTCATTTAGAACATTTAGATATTTTGGAGATTAAGTTAAGTCG 3487  
Qy 591 AACCGCGCGGAAGAGGCTGTCTGCCAGTTTGGCGGAATCCACCCGAGCAAAACAT 650  
Db TACAACGCTATCAAGCGAGATTTACCAGTTATTTCTGTGTACAGACGTCCTCAATTAATC 3427  
Qy 651 CGATGAGATCTTTAAAGCGCTGAAGCGCAGAC---TTACCCCATCTTTGTGAAGGCGAGT 707  
Db ATATGAATTTAGCAAAAGAAATTTGCAGAAAGAGCTGGTTTCCCGCTAATGATTAAGCCAC 3367  
Qy 708 TGCCTGCTGCGGACGCGGTATGCTGTTTGTGCTTTCCTGATGAGCTTCGCAAT 767  
Db AAGTGTGCGCGGTAAAGGTATGAGATGTTGCTGTAAGAAAGTGAATTTAGAAGATGC 3307  
Qy 768 AGCAACAGAGCATCTCTGTGAAGCTGAAGCGCTTTCGCGCATGCGCGGTATATGCGA 827

Db 3306 TTCCATAGAGCAAAATCAGAAGCTGAAATCAATTTGGTAATAGTGAAGTTTACATAGA 3247  
Qy 828 ACGTGTGTGATTAACCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGA 887  
Db 3246 AAGATACATTGATTAATCCAAAGCATATTGAAGTACAGTCAATAGTGCAGCATGGAAA 3187  
Qy 888 AGTTGTACACCTTTATGAACGTGACTGCTCACTCGACGGTCTGCACCAAAAAGTTGTGGA 947  
Db 3186 TATCGTACACTATTTTGAACGTGATTTGTTCAAGTCAACCGTCTCATCAAAAAGTTGTGGA 3127  
Qy 948 AATTGGCGCACACAGCATTTGGATCCAGAACTCGGTATCGCATTTGTGGCGATGCAGT 1007  
Db 3126 AGTTGCACCATCAGTTGATTAATCATCAATTAAGTCAAGTATTGTGATGCTGCAT 3067  
Qy 1008 AAAGTTCTGCGCTCCATTGTTTACAGGGGCGGGAACCGGTGAATTTCTTGGTGCATGA 1067  
Db 3066 TCAATTGATGGAATAATTAATAATGTCAATGCAAGTACTGTTGAATTTCTAGTATCTG- 3008  
Qy 1068 AAAGGGCAACCAAGCTTTCTGCGAAATGAACCCAGTATCCAGGTTGAGACACCGGTGAC 1127  
Db 3007 --GTGACGAATCTTCTTTATAGAAATTAACCTCGGTGACAGTACAGTACAAATTAC 2950  
Qy 1128 TGAAGAATCAACCGAGGTGGAACCTGGTGAAGGGCGAGTGGCTTGGGTGCTGTGCAAC 1187  
Db 2949 AGAGATGTTAAACAGGAATTGATATTGTTAAGACACAAATTTTAGTTCAGCAGGTGCCGA 2890  
Qy 1188 CTTGAAGGAATTGG-----GTCTGACCCCAAGATAAGATCAAGACCCACCGTGCAGC 1238  
Db 2889 TTTATTTGGTGAAGAGATTATATGCCCGCAACAAAGATATTACAATTAGGCTATGC 2830  
Qy 1239 ACTGCAAGTCCGCATCACACGGAAGATCCAAACAAACGGCTTCCGCCAGATACCGGAAC 1298  
Db 2829 CATCCAATGTGTAATTACAACAGAAATCCCGTTAAATGATTTTCATGCCGATACCTGAAC 2770  
Qy 1299 TATCACCGGTACCTCACCAGGCGGAGCTGGCTTGTCTTGAAGTGCAGTCAAGT 1358  
Db 2769 AATCAATTGCTATCTTCAAGCGGGGGCTTGGTGTACGTCTTGTGCTGGAGATGGTTT 2710  
Qy 1359 C---GGTGGCAAAATCACCGCACATTTGATCTCAATGCTGGTGAATGACCTCCCGTGG 1415  
Db 2709 CCAAGGTGCTGAGATACACCTTATTATGATTCATTACTCGTAAATTTATACACAGC 2650  
Qy 1416 TTCCGACTTTGAAACTGCTGTGCTGCGACAGCGCGGTGGTGGTGAATTCACCGTGC 1475  
Db 2649 GATATCAATTTAAACAGCAGAGAAATAATGGTACGCTCATACAGAAATTCGTTATTCG 2590  
Qy 1476 TGGTGTGCAACCAACATTTGTTCTTGGTGGCTTGTGCGGAAGAGGACTTCACATTC 1535  
Db 2589 TGGTGTAAAACATAATTTCCATTTCTTAATTAATGTAAGATATAAAGTTTCAACAG 2530  
Qy 1536 CAAAGCGCATGCGCACCGGATTCATTGCCGATCACCGCACCTCTTCAGGCTCCACCTGC 1595  
Db 2529 TGGTGATTACAACTAAATTTTGAAGAAACACCAAGAACTTTTCGACATTCAGCCGCTC 2470  
Qy 1596 TGAATGATGACGAGGAGCGATCCTGATTAATTTGGCAGATGTCACCGTGA-----C 1647  
Db 2469 TCTAGATAGAGTACTAAACATTAAGATATATAGTAATGTAACTTAATTTGGTTTCCC 2410  
Qy 1648 AAGCCTCATGTTGCGTCCAAAGGATGTTGAGCTCTTATGATGATGCTGCTTAACATC 1707  
Db 2409 AATGTTGAGAAACGTCGGAACCCAGACTATGAATTTAGCATCAATTTCCAACTGTATCTTC 2350  
Qy 1708 AAG-GATCTGCCACTGCCAGGTTTCCGTGACCGCTGAGAGGAGCTTGGGCCAGCGCG 1766  
Db 2349 AAGTAAATCGCTTCATTTAGTGTGACGAAACAAATTTGCTGTGATGATAGTGGTCCAAAGG 2290  
Qy 1767 GTTTGCTCGTGAATCTCCGTGAGCAGGACGACTGGCAGTTATGTATACCACTTCGCGGA 1826  
Db 2289 TGTAGCTGAATGGGTTTAAAGAGCAGGATGATGTCTTACTAAACAGATACAACTTTAGAGA 2230  
Qy 1827 TGCACACCACTCTTCTGCGACCGGAGTCCGCTCATTTGCACTGAACCTCGCGGAG 1886  
Db 2229 TGCACACCACTCATTTAGTGTACAGAGTTAGAACTAAGGATATGATTAATTCGATC 2170

Qy 1887 GGCCGTGCGAAGAGTGAATCCTGAGCTTTTCTCGTGGAGGCTGGGGGGGCGGACCTA 1946  
Db 2169 CAAAACAGCGGACGTAATTTAAAGATGGTTTCTCACTAGAAATGTGGGGGGTGTACAT 2110  
Qy 1947 CGATGTGGCGATGCGTTTCTCTTTGAGGATCCGTTGGGACAGGCTTCGACGAGCTGCGGA 2006  
Db 2109 TGAATGTGCATATAATTTCTTGAAGGAAACCCATGGGAAACGACTTGAACGCTACGTAA 2050  
Qy 2007 GCGGATGCGGAATGTAACATTTAGATGCTGTTCCGCGCGCAACACCGTGGGATACAC 2066  
Db 2049 AGCTATTCCAAATGTAATTTTCCAAATGTTGTAGCTCTTCAACGCGAGTTGGTTATAA 1990  
Qy 2067 CCGGTACCCAGACTCCGCTGCGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGA 2126  
Db 1989 AAACATCTCTGATATGTTTATTAATTCGTACAGAAAGTGTAAAGCAGGCATAGA 1930  
Qy 2127 CATCTTCGCAATCTTCGACGCGCTTAAAGCGTCTCCAGATGCGTCAGCAATCGAGC 2186  
Db 1929 TGTCTTTAGAAATTTTCGATTTCAATTAACCTGGTAGATCAAAATGAAAGTTGCCAATGAAGC 1870  
Qy 2187 AGTCTGAGACCAACACCGGCTAGCCGAGTGGCTATGGCTTATTTCTGGTGATCTCTC 2246  
Db 1869 AGTACAAAGAGCGGCAAAATCTCAGAGGTACTATTTGTTATACAGGTGACATTTTAA 1810  
Qy 2247 TGATCCAAATGAAAGCTCTACACCTCGATTAACCTTAAGATGGCAGAGGATCGT 2306  
Db 1809 TCTGAGCGATCAAAACATTTATCTTTAGATTAATGTCTAAACTAGCTAAGAGTTAGA 1750  
Qy 2307 CAAAGTCTGCGCTCACTCTTGGCCATTAAGGATATGGTGTCTGCTTGGCCAGCTGC 2366  
Db 1749 AGTGAAGTTTCCATATTTTAGCGANTTAAGATATGCGAGGCTTATTAAACCTTAAGC 1690  
Qy 2367 GGTAAACCAAGCTGGTCACCGCACTGCGCGTGAATTCGATTCGCGAGTGCAGTGCACAC 2426  
Db 1689 CGCTTACGAATTTGATGGTGAATTAATCAGCTGATGATTTACCAATTCATCTTCACAC 1630  
Qy 2427 CACGACACTGCGGTGGCCAGCTGCAACCTTCTTGTGCGAGCTCAAGCTGTGCGACA 2486  
Db 1629 TCATGATACAAGTGGTAATGGTTTATTAACATACAAAACAGCAATAGATGCTGGTCCA 1570  
Qy 2487 TGCTGTTGACGGTGTCTCCGCAACCTGCTGCGCACCTCCCGACCACTCCCGCTGCTGC 2546  
Db 1569 TATCATTTGATCTGCTGTTGCTTCAATGATGTTGTTTAAACAGTCAAGCAAGCGCAATTC 1510  
Qy 2547 CATTTGCTGCAATTCGCGCACACCGCTCGCATACCGGTTGAGGCTCGAGGCTGTTTC 2606  
Db 1509 GTTATATTATGCAATTAATGGCTTCCCGCCCACTTGAACCTGATGATGGA 1450  
Qy 2607 TGACCTCGAGCGCTACTGCGGAGCAGTGGCGGAGCTGTACTCTGCCATTTGAGTCTGGAAC 2666  
Db 1449 GTCACTTAGTCAATTTGGTCAACTGTAAGTCTTATTTATTCAGACTTTTGAAGTGTAT 1390  
Qy 2667 CCCAGGCCCAACCGGTGCGGTCTACCGCAACGAAATCCCGAGGCGGACAGTTGTCCAACCT 2726  
Db 1389 CAAATCACCAGTACTGAAATTTTATCAACATGAATGCGCTGGTGGACAGTATTCGAATTT 1330  
Qy 2727 GCGTGCACAGCGCCACCGCACTGGCGCTTGGGATCGTTTGGAACTCATCGAAGCAACTA 2786  
Db 1329 AAGTCAACAGCTTAAAGTTTAGTTTAGCGGAAAGATTGATGAAGTCAAGATATGA 1270  
Qy 2787 CGCAGCGGTTAATGAGATGCTGGGACCGCCCAACCAAGGTCACCCCATCTCCAGGTTGT 2846  
Db 1269 TCGCAGAGTGAATTTCTTATTTGTTGATATCGTAAAGTAACCACTCGTCTAAAGTAT 1210  
Qy 2847 TGGGCACTCGCACTCCACCTCGTTGTTGGGGGTGGATCCAGCAGACTTTGCTGCCGA 2906  
Db 1209 TGGTGAATGCGACTTTTATGTTACAAATGATCTTGTGAACAACTCCGTGATTACAGA 1150  
Qy 2907 TCCAACAAAGTACGACATCCAGACTCTGTCTATGCGGTTCTCGCGCGGCGAGCTTGTGTA 2966  
Db 1149 TGGCTATAAATTAGATTTCCAGATCAGTAGTGTGCTCTTCTTCAAGGTTGAAATAGGACA 1090

QY 2967 CCTCCAGTGGTGGCCAGACCCACTCGCGCACCCGCGACCTGGAGGCCCTCGAAGG 3026  
Db 1089 ACTGTAAATGGTTTAAATAAGATTTACAGCGGTTATTTTAAAGGCCAAGCACT 1030  
QY 3027 CAAGGCACCTCTGACGGAGTTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 3086  
Db 1029 AACAGCTCGTCCAGGTGAATATCTAGAGCCAGTTGATTTTGAAGAAGTCCGTGAGTTGCT 970  
QY 3087 CAAGGAACGTGCAATAGCC-----TCACCCGCTGCTGTTTCCC 3125  
Db 969 TGAAGAAGAGCAACAGGTCCTGTTACGGAGCAAGATATTATTAGTTAGTACTATATCC 910  
QY 3126 GAAGCCAAACGAGAGTTCTCTCGAGCACCGTCCGCTCGCGACACCTCTGCGCTGA 3185  
Db 909 AAAAGTATATGAACAATAATTAACAATAAATAAATCAATACGGAACCTTATCGTTACTTGA 850  
QY 3186 TGATCGTGAATCTTCTACGGCTGCTCGAAGCGCGAGACTTTGATCCGCTGCCAGA 3245  
Db 849 TAGCGCTACATCTCTTTGGAATGCGTAATG---GTGAACAGTAGAATAAGAAATCGA 793  
QY 3246 TGTGCGCACCCACTGCTTTGTTGCGCTGATCGATCTCTGAGCCAGAGATAGGGTAT 3305  
Db 792 TAAAGGTAAACGATTAAATTAATACTAGAAACGATTAGTAACACAGATGAATGGTAA 733  
QY 3306 GCCCAATGTTGTGGCCAAAGTCAACGCGCCAGATCCGCCCAATCGGTGCGTACCGCTC 3365  
Db 732 TAGAACGATTACTATGCGATGAATGTTCAAGCGAGCTATTACATTAAAGATGAAAA 673  
QY 3366 CGTTGAGTCTGTCCAGCAACCGCAGAAAGGAGATTTCTCCACAGGGCCATGTCG 3425  
Db 672 TGTGCTATACAAATGCGAAGTTAAGCCAAAGCAGATAGATATCCAACTCATATCGG 613  
QY 3426 TGACACCATTCGCTGGTGTGTGTAC---CGTGAATGTTGCTGAAGGTGATGAGGTCAAGGC 3482  
Db 612 TGCTCAATCCAGGTTCACTAGTCAAGTTAGTGTAGGTGAAGTGTGAAGC 553  
QY 3483 TGAAGTGAAGTGGCAATCATCGAGGTATGAGATGGAAGCAACATCACTGCTTCTGT 3542  
Db 552 TAATCAAGCCGTTGCTAATTACTGAAGCTATGAAATGGAATGGAACCAACAAATCAAGACCAAT 493  
QY 3543 TGACGGCAAAATCGATCGCTGTTGTTCTGCTGCAACGAGGTGGAAGTGGGACATT 3602  
Db 492 TGACGGTGTGATTAAACAGTACTGTAATATGTTGACACATAGCAGAGCGGATTT 433  
QY 3603 GATGCTGCTGCTT 3615  
Db 432 ATTAATCGAAATT 420

RESULT 9  
US-09-107-532A-1156  
; Sequence 1156, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 1156:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2922 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURES:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...2922  
SEQUENCE DESCRIPTION: SEQ ID NO: 1156:  
US-09-107-532A-1156  
Query Match 13.2%; Score 476.2; DB 4; Length 2922;  
Best Local Similarity 49.9%; Pred. No 2.5e-125;  
Matches 1425; Conservative 0; Mismatches 1408; Indels 24; Gaps 8;  
QY 199 GTGTGCACTCACCATCTTCAACGCTTCCAGCATTCACCGCTCTTTTCTGCTGAGTGTCCGC 258  
Db 1 GTTTATATAAAAAAGATCTAGGAGTGGTTCGATGAAAAAGTTTTAGTTGTTAAACAGG 60  
QY 259 GGCGAATTCGGTCCGCTGCTTCCGTGCAAGCTCGAAACCGTGCAGCCAGGTAGCT 318  
Db 61 GGTGAATTCGCTGACGTGATTTCTGCTGTTGACAGATTAGGAATTAACATGTTGGA 120  
QY 319 ATTTACCCCGCTGAAGATCGGGGATCATTCACCGCTCTTTTCTGCTGAGTGTCCGC 378  
Db 121 ATCTATGCAAAAAAATGAATATTCGTACACCGTTTCAAAGCGGATGAAGCATATTTA 180  
QY 379 ATTGTCAGGAGCTCACCAGTCAAGCGTACCTGACATCGATGAAATATTCGTTGCA 438  
Db 181 GTAGCAACAGAAAAACCGATTGATCGGTATCTAGATTGAAGGATCATTTCTATC 240  
QY 439 GCTAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACCGCTTCTGCTGAAAAATGCC 498  
Db 241 CGGAAGATCGCGAGCTGACGCTATTTCATCCGGTTTACGGTCTTTTGTGGAACCTTG 300  
QY 499 CAGCTTCCCGGAGTGTGGGAAACGGCATTTCTTTTATTTGCCCCCAACCCAGAGTT 558  
Db 301 AATTTCGCAACGCTGCGAAGAAAAAGGCATCATTTTGTGGACCCCAACTGCATCAT 360  
QY 559 CTGATCTCACCGGTGATAGTCTCGCGCGGTATACCGCCGCGTACCGCCGCGAAGAGGCTGG---TCTG 615  
Db 361 TTAGATATTTTGGAGATATAATCAAGCAAAAGCTCAGCTATCGAAGCGGGATGCT 420  
QY 616 CAGTTTTGGCGGAATCCACCCGAGCAAAACATCATGATGATCGTTAAAGCGCTGAA 675  
Db 421 TCGATTCCAGGCACAGATGGACCAATCGTTTCTATAGACGATGCTTTGGAGTTTGGGAAA 480  
QY 676 GGCCAGACTTACCCCATCTTTGTGAAGGCACTTCCCGTGTGGTGGCGGACGCGGTATCGGT 735  
Db 481 CAATACGTTTATCGATTATGATCAAGGACCTCTATGCGGCGCGGACGCGCATGCT 540  
QY 736 TTGTGTTGTTCCCTGATGAGCTTCGCAAAATAGCAACAGAGCATCTCTGTAAGCTGAA 795  
Db 541 GTAGCTCAGATGAAAGAGTCAAGAGAGGCTACGAACTGCGGAAAGCGAAGCAAAA 600

QY 796 GCGCTTTGCGCGATGGCGGATATATGTGGAACGTGCTGTGATTAACCCCTGAGCATATT 855  
DB 601 GCAGCCTTTGGTAGTGAAGTGTATGTTGAAAAATACATCGCTAATCTTAAGCATATC 660  
QY 856 GAAGTGACAGATCCTTTGGCGATACACATGGGAAGTGTGACACCTTTATGAACGTGACTGC 915  
DB 661 GAAGTACAGATTTCTCGGGATACGCATGGAAACGTCACTCCATCTTTTGAAGAGATTGT 720  
QY 916 TCACGTGAGCTGTGTCACCAAAAAAGTTGTGGAATTTGGCCGACGACAGCATTTGGATCCA 975  
DB 721 TCTGTTACGCGCGCTGACCAAAAAAGTAGTAGAAGTAGCGCTTGTGTCTCTATGAATGAA 780  
QY 976 GAACTGCGTGTGTCATTTGTGCGGATGTCAGTAAAGTTCTGCCCTCCATTGGTTACCG 1035  
DB 781 CAGCAACGTCAAAAACTGCGGACGAGCTGTCCCATTTGATGAATATGTCGGGTATGTC 840  
QY 1036 GCGCGGGAACCGTGGAAATCTTGGTGTGATGATAAAGGGAACCAACGCTCTTCAATCGAAATG 1095  
DB 841 AACGCACGTACAGTAGAATTTCTTTGTCGAAGGAGATG---ATCTTTACTTTATCGAAGTC 897  
QY 1096 AACCCACGTATCCAGTTGACACACCGTGTGATGAGAAAGTCAACGAGGTGGACCTGGT 1155  
DB 898 AACCTTCGCTCCAAAGTGGAAACATACGATCAAGAAATGATACAGATATTGATATCGTA 957  
QY 1156 AAGCGCAGATTCGCTTGGCTGTGGTGTGCAACCTTG---AAGGAATGGGCTGAC---C 1209  
DB 958 ACAACGCAACTGTTGATTTGCTCAAGGTTGGATCTGCATAAAGAAATCGGATTCGCTCAG 1017  
QY 1210 CAAGATAAGATCAGACCCAGGTGACACATGCGATGCGCATCACACGGAAGATCCA 1269  
DB 1018 CAAGAAGAAATTAATTAATTTGTTTTCAGCCATTTAATGCCGAATCACTACGAAGACCG 1077  
QY 1270 AACACGGCTTCCGCCAGATACCGGAATATCAACCGGTACCGCTCAACCGAGCGAGCT 1329  
DB 1078 TTGAATAACTTTTACCAGACACAGGGAATTCGATACGTATCTGTACCTGGTGGTTTT 1137  
QY 1330 GCGTTCGCTTGTGACGTTGACGCTCAGCTCAGCTCGGTGGCA---AATCAGCGACACTTTGAC 1386  
DB 1138 GCGCTACGTGTGATGTGCGAAATGCTTACCGGCTATGTAGTAGACACCGTATTTTGT 1197  
QY 1387 TCCATGCTGTGAAATGACCTGCGGTGTTCCGACTTTGAACTGTGTTGCTCGTCA 1446  
DB 1198 TCATTGCTAGTAAAGTATGACGATGTCATGCTACCTTTGAAACAGCCATTTCAGAAATG 1257  
QY 1447 CAGCGCGGTGGCTGAGTTCACCGTGTGCTGTGTGTCACCAACCATGTTCTTGGT 1506  
DB 1258 GAACGTTGCTTGAAGAATTTAGAAATCCGTGGAGTGAATAAACAATTTCCATTTATGCTC 1317  
QY 1507 GCGTTGCTGCGGGAAGAGACTTCACITTCGAAGCGCATCGCCAGATTTCATTGCGGAT 1566  
DB 1318 AATGTATACACATCTCTGMAATTCATCAGGAATGCAAAAACAACCTTCATTGATAGT 1377  
QY 1567 CACCGCACCTCCTTCAGGCTCCACTGTGATGATGAGCAGGAGCGCATCCTGGATTAC 1626  
DB 1378 ACAGCAACATTTATTTGAATTTCTCGACTACGAGACCGCGGAAACAAAACGATGAAATAT 1437  
QY 1627 TTGGCAGATGTCAACCGTGAACAAGCTCATGTGT-----CGTCCAAAGGATGTTCAG 1681  
DB 1438 ATCGGTGAAATTTACTGTGAAATGGTTTTTCAGGTATCGAAAGTGGGGAACCAATTTAT 1497  
QY 1692 CTCCTATCGATAGCTGCTTAACATCAAGGATCTGCCACTGCCACGCGGTTCCGCTGACC 1741  
DB 1498 GAAGAACCAACGATATGCTTAAAGATCTAATCAAGAACTGATTACGTAACGCGCTAAGAAT 1557  
QY 1742 GCCTGA-AGCAGCTTGGCCCGACGCGGTTGCTCGTGTATCTCCGTGACGAGCGCATG 1800  
DB 1558 GTAATTAGTGCAGAGGTGCCGATGCTTTGGTGTGAAATGATCAAGGACAGGAAATTTG 1617  
QY 1801 GCAGTTACTGTATACCACTTCCGATGACACACAGTCTTTGCTTGGCAACCGAGTCCG 1860  
DB 1618 TTGTTGACCGATACGACATTTAGAACGCCCATCAAGTCTGCTAGCAACTCGTGTCCG 1677  
QY 1861 TCAATTCGCACTGAGCTCTCGCGCAGAGCGCGTCCGAAAGCTGACTCTCTGAGCTTTTGCC 1920

DB 1678 ACAAAGGATTTCAACGAGATTCGACGTTTGACTGGAGAGGCTCTCTCTGAACTCTTTCA 1737  
QY 1921 GTGAGGCTTGGGCGCGCGACCTACGATGTGGGATGCGTTTCTCTTTGAGATCCG 1980  
DB 1738 AGTGAGATGTGGGGGAGCAACATTTGATGTGGCTTACCGCTTTTGAATGAAGATCCT 1797  
QY 1981 TGGGACAGCTCGACAGCTCGCGAGCGATGCCGAATGTAACATTCAGATGCTGCTT 2040  
DB 1798 TGGCAACGGTTGAGAAAAATTCGTTCTTTAATGCTTAATAGTTGCTGCAATGCTTTTC 1857  
QY 2041 CCGCGCCGCAACACCGTGGATACACCCGCTACCCGATCCGCTGCGCGCGTTGTT 2100  
DB 1858 AGAGGTTCAATGCGGTTGGGTATTTCTAATTTATCTGCAATGTATTGGTGGATTTGTC 1917  
QY 2101 AAGGAAGCTGCAGCTCCGCGGTGGACATCTTCGCACTCTTCGACGCGCTTAACGACGTC 2160  
DB 1918 AAGGAAGCAGCAGCGCAAGGAATCGATGTTATTCGAATCTTTGATAGTTGAACTGGACT 1977  
QY 2161 TCCGAGATGCGTCCAGCAATCGACGAGTCCCTGGAGACCAACACCGCGGTAGCCGAGTG 2220  
DB 1978 CCACAAATGGAATAAAGTATCCAGCTGTACGGGATCTGGTA---AAATCGCAGAGCA 2034  
QY 2221 GCTATGGCTTATCTGGTGATCTCTCTGATCCAAATGAAAGCTGTACACCTGATTTAC 2280  
DB 2035 GCTATTTGTATATCTAGTGGGATATCAACGATCCAGTCCGGGCAAAATACATGTTCAGTAT 2094  
QY 2281 TACCTAAAGATGGCAGAGAGATCGTCAAGTCTGCGCTCAATCTTGGCCCAATTAAGAT 2340  
DB 2095 TATAAGACATGCAAAAGGAACTAGAAACAACGTGGTGGCATATCATTTGCAATCAAGAT 2154  
QY 2341 ATGGCTGTCTGTTGCGCCAGCTGGGTAAACCAAGCTGGTCAACGCACTGCGCGTGA 2400  
DB 2155 ATGGCTGGCTCTTAAACCGCAGGACGCTTATCGTTTGAATAGTGAATTTGAAGAAAG 2214  
QY 2401 TTGCACTGCCAGTGCAGTGCACACCAACGACATCGCGGTGGCCAGCTGGCAACCTTAC 2460  
DB 2215 ACAGATCTGCCATCCATCTGCACTACTCATGATACACGCGGAATGGAATCATCACTTAT 2274  
QY 2461 TTTGCTGAGCTCAAGCTGGTGCAGATGCTGTGAGGCTGCTTCGACCACTGTCTGCG 2520  
DB 2275 TCTGTGCTTCAAAAGCAGGAGTAGACATCTGTGAGCTGGCAATGAGTGAATGAGTGG 2334  
QY 2521 ACCACCTCCAGCATCCCTGCTGCTGCAATTTGTGCAATTCGCGCACACCCGTCGCGAT 2580  
DB 2335 AATACTAGTCAGCTAGCATGAGCAGCTTGTATTTATGCACTGCTCAATGGTCCACGCTG 2394  
QY 2581 ACCGTTTGAGCTCGAGCTGTTTCTGACCTCGAGCGGTACTGGGAAGCAGTGGCGGA 2640  
DB 2395 CCAGAAATCAGATTGAAATGCACAGAAATTAACCAATTTAGGGAAGATGTGGCGATG 2454  
QY 2641 CTGTACTCTGCCATTTGAGTCTGGAACCCCGAGCCCAACCGGTGCGGTCTACCGCCACGAA 2700  
DB 2455 TATTACAAACCGTTTGAACACCGCTTGAATGCTCTCTGAAACAGAGATATATATGATGAA 2514  
QY 2701 ATCCAGGCGGACAGTTGTCCAAACCTGCTGACAGGCGCACCGCACTGGGCTTGGGAT 2760  
DB 2515 ATGCCAGGAGGACAGTATTTCGAATCTACAGCAACAGGCAAAAGCTGTTCGGAATAGTCA 2574  
QY 2761 CGTTTCGAATCATCGAAGCAACTACGACGCGTTAATGAGATCTGGAGACGCCCAACC 2820  
DB 2575 CGCTGGGATGAAATCAAAACAAATGATCATATGTCACCTTGTATGTTTGGCGATTCGTT 2634  
QY 2821 AAGGTCAACCCCATCTCTCAAGGTTGTGGCGACTTCGCACTCCACTCGTCTGGTGGGCT 2880  
DB 2635 AAAGTAACCTCCATCATCTAAAGTAGTTGGAGATATGGCACTATTATTGTTCCAAATGAT 2694  
QY 2881 GTGGATCCAGCAGACTTCTGCTGCCGATCCACAAAGTACGACATCCAGACTCTGTCATC 2940  
DB 2695 CTTACAGAGAAGATATTTATGAAAAGCGGAAACACTAAGTTTCCCTGAACTGTGTG 2754  
QY 2941 GCCTTCTCGCGCGCAGCTTGTGTAACCCCTCCAGTGGCTGGCGCAGAGCACTGCGCAC 3000

Db 2755 ACTTCTTCCAGAGAACTAGGCAACCCGTTGGCGGATTCCTCTAAATAATTCAGAAAG 2814

Qy 3001 CCGGCACATGGAAGCGCGCTCCGAGAGGCAAGCACCTC 3037

Db 2815 ATTATTTCTTAAGAGCAGACCTGCATTGAATGAACGTC 2851

RESULT 10

US-09-543-681A-1699

Sequence 1699, Application US/09543681A

Patent No. 6605709

GENERAL INFORMATION:

APPLICANT: GARY BRETON

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

FILE REFERENCE: 2709, 1002-001

CURRENT APPLICATION NUMBER: US/09/543, 681A

CURRENT FILING DATE: 2000-04-05

PRIOR APPLICATION NUMBER: US 60/128,706

PRIOR FILING DATE: 1999-04-09

NUMBER OF SEQ ID NOS: 8344

SEQ ID NO 1699

LENGTH: 1362

TYPE: DNA

ORGANISM: Proteus mirabilis

US-09-543-681A-1699

Query Match 8.7%; Score 313.4; DB 4; Length 1362;

Best Local Similarity 55.4%; Pred. No. 4.4e-79;

Matches 721; Conservative 0; Mismatches 556; Indels 24; Gaps 5;

Qy 236 AAAAGATCTTGTAGCAACCGCGGGAATCGCGTTCGGTGCAGCACTCG 295

Db 20 AAAAATCTCTATTGCAACCGGTGGTGAATTTGCACTGGTATCTTAAGG 79

Qy 296 AAACCGGTGAGCAACCGGTAGCTATTACCCCGTGAAGATCGGGATCATTTCCACCGCT 355

Db 80 AACTTGGATCAAGACAGTCGCGTTCACTCCAGCGACACCGTGAATTTAAACACAGTTC 139

Qy 356 CTTTTCCTCTGAGCTGTCGATTTGTTACCGAGGCTCACAGTCAAGGCGTACCTGG 415

Db 140 TGTGTCAGACGAGACTATCTGTATTGTTGCGCGTCTTCAGCA---AAAGTTACTTAA 196

Qy 416 ACATCATGAAATATTCGGTGCAGCTTAAAGATTTAAAGCAGATGCCATTTACCCGGAT 475

Db 197 ATATTCCGCAATTATTCGCGGCGAGAGATAAGTGGCGCGCAAGCAATTCACCCAGAT 256

Qy 476 ACGCTTCTCTGTAATGCGCGCTTCCCGCGAGTGTGCGGAAACGCAATTTT 535

Db 257 ATGCTTCTCTGTAATGCGCAATTTTCCGCAACAGTTGAACGCTCAGGCTTTATT 316

Qy 536 TTATTGGCCCAACCCAGAGTTCTTGTATCTACCGGTGATAAGTCTCGCGGTAACCG 595

Db 317 TTATTGGCCCTTAAGCGGAACCAATTCGCTTAATGGTGTAAAGTTTCCGCTATTGAAG 376

Qy 596 CCGCGAAGAGGCTGTGTCTGCAAGTTTGGCGAATC-----CACCCGACAAACAA 649

Db 377 CGATGAAAGAGCGGTGTCTTGTGTACAGGCTCAGATGGCCCAATAGTAAAGATA 436

Qy 650 TCGATGAGATCGTTAAAGCGTGAAGCGCAGACTTACCCCATCTTTGTGAAGGAGTTG 709

Db 437 CAGCAAAATATTCGAATTCGCAACGCAATTTGGTTACCTGTATCATCAAGCATCAG 496

Qy 710 CCGGTGTGCGGACCGGCTATCGTTTGTGTCTTCACTGATGAGCTTTCGCAATTAG 769

Db 497 GTGTGCGGCTGTGCGGCTATCGTTTGTGTCTCAGAAAAAGACTTTAGCGCAAGCAA 556

Qy 770 CAACAGAGCATCTCGTGAAGCTGAAGCGCTTTTCGCGATGCGCGGTATATGTGCAAC 829

Db 557 TCTCATGACCGGTGCGGAGCCAAAGCGCAATTTAGCAAGATATGTTCTATATGAAA 616

Qy 830 GTGCTGTGATTAACCTTCAGCATATTGAAGTGCAGATCTTGGCGATCACTTGAAGAG 889

Db 617 AGTACCTTGAAATCCACGCCACGTCGAAATTCAGGTGATGGCCGATGCAAGTAATG 676

Qy 890 TTGTACACCTTTATGAACGTGACTGCTCACTCAGCGTCTCACCACAAAAGTTGTGAAA 949

Db 677 CTATCTATTAGTGAACGTGACTGCTCAATCAACAGTCGCGCACCAAAAAGTTGAG 736

Qy 950 TTGCGCCACACAGCAATTTGGATCCAGAACTCGGTGATCGCAATTTGCGGATGAGTAA 1009

Db 737 AAGCACACGACCGGGTATTACCCCTGAGATCCGTAATAATATCGGTGAACGCTGTGCAA 796

Qy 1010 AGTTCTGCGCTCCATTGTTTACAGGGCGCGGGAACCGTGAATTTCTTGTGCGATGAAA 1069

Db 797 ATCCCTGTAATGAATTTGCTACCGCGGTGCGGTACGTTTGAATTC---CTCTATGAAA 853

Qy 1070 AGGCAACACAGCTTTTCATCGAAATGAACCCACGATCCAGTTGAGACACACCGTACTG 1129

Db 854 ATGCGAATTTCTACTTTATCGAAATGAATACCCGTTATTCAGGTTCAGCATCTCTACTG 913

Qy 1130 AAGAAGTCCAGGAGTGGACCTGGTGAAGCGCAGATGCGTTGGTGTGCTGTGCAACCT 1189

Db 914 AGATGATCACCGGTGTTGACCTTTATCAAGAGCAACTGCGTATTGATCAGGCTTACC-- 971

Qy 1190 TGAAGGAATTTGGTCTGACCCCAAGATPAAGATCAAGACCCACGCGTGCAGTGCCT 1249

Db 972 -----ATTATCAGTCACGCAAGATCAAAATTCAGTTTCATGACATGCTATTGAGTGC 1024

Qy 1250 GATCACCAAGAGATCCAAACACCGCTTCGCCCCAGATACCGGAACATATCACCGCT 1309

Db 1025 GTATCAACGCAAGATTCACAAA---CTTTCTGCAAGCCCGGAAACCATCATCTCTGTT 1081

Qy 1310 ACCGCTCACCGCGGAGCTGGCGTTTCTGTTTGAACGGTGCAGTCAAGTTCGGTGGGAAA 1369

Db 1082 TCCACTCACCGCGGATTTGGTGTACGTTGGGAATCATATTTACGCAAGTTTACCGG 1141

Qy 1370 TCACCGCACATTTGACTCCATCTGTTGAATGACCTGCGGTGTTCCGACTTTGAAA 1429

Db 1142 TTCACACACATGATTCATGATTGTTAAATTTGATCACTTACGTTGAAAACAGTGA 1201

Qy 1430 CTGCTGTGCTGTCACACGCGCGCTTGGCTGAGTTTCCCGTGTCTGGTGTGCAACCA 1489

Db 1202 TTGGATTTCTGATGAAATGCGTTGGCAGAACTTATTTAGCGGCAATTAAACCA 1261

Qy 1490 ACATGTTTCTGCGTGGCTGCTCGCGGAGAGGACTTC 1530

Db 1262 ATATTGAGTACACCAATTCATTGATGATGAGCAATTC 1302

RESULT 11

US-07-956-700B-5

Sequence 5, Application US/07956700B

Patent No. 5539092

GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki

TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA

TITLE OF INVENTION: Carboxylase

NUMBER OF SEQUENCES: 116

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: 321 No. 5539092th Clark Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/956,700B

FILING DATE: 19921002

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Thomas E. No. 5539092thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1362 base pairs  
TYPE: Nucleic acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Oligonucleotide  
US-07-956-700B-5

Query Match 8.68; Score 313; DB 1; Length 1362;  
Best Local Similarity 53.98; Pred. No. 5.7e-79; Mismatches 600; Indels 15; Gaps 3;  
Matches 718; Conservative 0;  
Db 232 TTCAAAAGATCTTGGTAGCAACCGCGGGAATCCCGTCCGTTCCGTGCGAGCA 291  
7 TTCAACAGATCTGATCGCAATCGCGGGAATCCCGTCCGCAATCTCCGCACTTGT 66  
Qy 292 CTCGAAACCGGTGACGACCGAGTATTTACCCCGTGAAGATGGGATCATTCAC 351  
Db 67 CAAGAATCTGGGATCGGACGATCGCGTTCACCTCACTGTGGATCGAAGCGCTCCAT 126  
Qy 352 CGCTCTTTTGTCTGAACTGTCCGATTTGGTACGAAGGCTCACCAGTCAAGGCGTAC 411  
Db 127 GTGCAGTTAGCGGACGAAGCGTCTGATTGG---CGAAGCGGACGACCAAAAGCTAT 183  
Qy 412 CTGACATCGATGAATATTCGGTGCAGCTTAAAGATTTAAGCAGATGCCATTTACCG 471  
Db 184 CTCGAATATCCCAACATCATTTGGCGGCGCTGACCCCTAATGCCAGCGCATTCACCCC 243  
Qy 472 GGATACGGCTTCCTGTCTGAATAATGCCAGTTCGCCCGGAGTGTGGGAAACGGCAAT 531  
Db 244 GGCTATGGCTTTTGGGAGATGCGCGCTTTCAGAAATCTGGCGCATCACCATCTC 303  
Qy 532 ACTTTTATGGCCCAACCCGAGAGTTCATCTCAACGGTGTATAGTCTCGGC---G 588  
Db 304 ACCTTTATGGCCCGAGCCCGATTCGATTCGAGCCATGGCGGATAAATCCACGCTAAG 363  
Qy 589 GTAAACCGCGGAGAGAGCGTGTCTGCCAGTTTGGCGAATCCACCCCGAGCAAAAC 648  
Db 364 GAACATGACAGCGGTTCGGTTCGAGATTCGGCGAGTGAAGTCTGCTGACGGAT 423  
Qy 649 ATCGATGAGATCGTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAGGCGATT 708  
Db 424 GTTGATTCGGCTGCCAAAGTTGTGCCGAGATCGGCTATCCCGTCAATGATCAAGCGAG 483  
Qy 709 GCGGTGTGGCGGAGCGGTATGCTTTGCTTCACTGATGAGTTCGCAAAATTA 768  
Db 484 GCGGGGCGGTGTGCGCGGTATGCGGTGTGTGAGTACCTGAGATCTGGAATACTG 543  
Qy 769 GCAACAGAGCATCTCTGAAGTGAAGCGGCTTTCGCGAGTGGCGGATATATGTCGAA 828  
Db 544 TTCCTTGTGCCCCAGGAGAGCGGAGCGCTTTGGGAATCAGGACTGTATCTCGAA 603  
Qy 829 CGTGTGTGATTAACCTTCAGATTTGAAGTGCAGATCTTGGCGATCAGCTGGAGAA 888  
Db 604 AAATTTATCGATCGGCCACGCACTTGAATTCAGATCTTGGCGGATGCTACGGCAAT 663  
Qy 889 GTTGTACACCTTTATGACGTGCTCCTGACGAGTGTCTCAACCAAAAGTTGTGCAA 948  
Db 664 GTAGTGCATCTAGGAGCGGATGCTCCATTCAGCTGTCAACCAAAAGCTGCTCGAA 723  
Qy 949 ATTGGCCAGACACATTTGATTCAGATTCAGATTCGATTCGATTCGATTCGATTCGAT 1008  
Db 724 GAAGCCCCAGTCCCGCGCTATCGGACAGCTTCGCGGAGAAATGCGGATGCGCGCTC 783  
Qy 1009 AAGTTCTCGCGCTCCATTTGTTTACAGGCGCGGAGACCGTGGAAATCTTCTGTCGATGAA 1068

Db 784 AAGTGTCTCAAGCGATCGGCTACATCGGTCCCGCACCGTGGAGTTTCTGGTCGATGCG 843  
Qy 1069 AAGGCGACACCGTCTTCATCGAATGAACCCAGTATCCAGGTTGAGCACACCGTCACT 1128  
Db 844 ACCGCGACTTCTTACTTCATGAGATGAATACCGCATCAAGTCGAGCATCCAGTCA 903  
Qy 1129 GAAGAAGTCAACGAGGTGGACCTGTGAAGCGGAGATGGCTTGGTGTGTGTGCAACC 1188  
Db 904 GAAATGATTACGGGACTGGACTTGAATCGGAGCAGATTCGGATTGCCAAGGCG---- 958  
Qy 1189 TTGAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACCCACCGTGCAGCACTGCACTGC 1248  
Db 959 ----AAGCGTTCGCTTCCGCAAGCGATATTCATCTGCGGCGCATGCGATCGAATGC 1014  
Qy 1249 CGCATCAACCGAAGATCCAAACAAAGCGTTCGCCCCAGATACCGGAATATCACCGCG 1308  
Db 1015 CGTATCAATCGGAAGATCCGGAATACAAATTCGCGCGGAATCTCGCGCATTACAGGC 1074  
Qy 1309 TACCGCTCACAGGCGGAGTGGGTTGCTTGAAGCGTGCAGCTCAGCTCGGTGGCGAA 1368  
Db 1075 TATTTACCGCGCGCGCGCGCTTGGTGTGCTGATTCGCCATGTTTATACGACTACGAA 1134  
Qy 1369 ATCAGCGCACACTTTGACTCCATGCTGTGAAATGACCTGCGGTGTTCCGACTTTGAA 1428  
Db 1135 ATTCCGCCCTATTACGATTCGCTGATTGGCAAAATTGATTGCTGGGTGCAACACGGGAA 1194  
Qy 1429 ACTGCTGTGCTGTGCAAGCGCGGTGCTGCTGAGTTACCGTGTCTGTGTGTGCAACC 1488  
Db 1195 GAGGCGATCGCGCGATGACGCTCTCTCGGGAATGCGGCATCACGGGTTCGCCAGC 1254  
Qy 1489 AACATTGTTTCTTCCGTCGTTGCTGGGGAAGAGGACTTCACTTCCAGCGCATCGCC 1548  
Db 1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCTGAGTTCCTGCGCGGGAATCTAT 1314  
Qy 1549 ACCGATTCATG 1561  
Db 1315 ACCAATTTGTG 1327

## RESULT 12

US-08-476-537-5  
Sequence 5, Application US/08476537  
Patent No. 5756290  
GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
TITLE OF INVENTION: Carboxylase  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5756290th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,537  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
FILING DATE: 10/21/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5756290thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090

TELEFAX: 1-312-755-4489

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1362 base pairs

TYPE: Nucleic acid

STRANDEDNESS: Single

TOPOLOGY: Linear

MOLECULE TYPE: Oligonucleotide

US-08-476-537-5

Query Match

Best Local Similarity 53.9%; Pred. No. 5.7e-79; Length 1362;

Matches 718; Conservative 0; Mismatches 600; Indels 15; Gaps 3;

```
QY 232 TTCAAAAGATCTTGTTAGCAACCGCGCGGAAATCGCGGTCCGTCTTTCGGTCAGCA 291
Db |||
QY 292 CTCGAACCGTGCACGACGATGATTTACCCCGGTGAAGATCGGGGATCATTCAC 351
Db |||
QY 67 CAAGAACTCGGGATCGGCACGATCGCGGTTCATCCACTGTGGATCGCAACGCGTCCAT 126
Db |||
QY 352 CGCTCTTTTCTCTCAAGCTGTCGCATTTGGTACCGAAGGCTCACAGTCAAGCGGTAC 411
Db |||
QY 127 GTCAGTTCAGCGACGAGCGGTCTGTATTGG---CGAAGCGCGCAGCAGCAAAAGCTAT 183
Db |||
QY 412 CTCGAATCATGAATTAATTCGATGAGCTAAAGTAAAGAGATGATTCATTCAGCG 471
Db |||
QY 184 CTCATATATCCCAACATCATTTGGCGGCGCCCTGACCCCTAAATGCCAGCGCCATTCACCCC 243
Db |||
QY 472 GGATAGCGTTCCTGTCTGAATGCGCAGTTCGCGCGAGTGTGCGGAAACCGCATT 531
Db |||
QY 244 GGTATGGCTTCTTGGCGGAGATGCGCGTTTGGAGAAATCTGCGCGATCACCATCTC 303
QY 532 ACTTTTATTTGGCCCAACCCCGAGGTTCTTGATCTCACCGGTGATAGTCTCGCG---G 588
Db |||
QY 304 ACCTTTATTTGGCCCGACCGCCGATTCGATTCGAGCCATGGCGGATAAATCCACCGCTAAG 363
QY 589 GTAACCGCCCGAAGAGGCTGTCTGCCAGTTTTCGCGGAATCCACCCCGAGCAAAAC 648
Db |||
QY 364 GAACAATGAGCGGTGCGGTTCGACGATTCGCGGAGTACGCGTCTGTGACGAT 423
QY 649 ATCGATGAGATCGTTAAAGCGTGAAGCGCAGATTAACCCATCTTTGTGAAGGAGTT 708
Db |||
QY 424 GTTGATTCGGCTGCAAGTTGTGCGAGATCGGCTATCCGCTCATGATCAAGCGAG 483
QY 709 GCGGTGTGCGGAGCGGTATGCTTTGTTGTTCTCACTGATGAGTTCGCAATTA 768
Db |||
QY 484 GCGGGGGCGGTGTCGCGGTATGCGCTGTGCGGTGACCCCTGAGATCTGGAAATCTG 543
QY 769 GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGCGATGCGCGGTATATGTCGAA 828
Db |||
QY 544 TTCCTTGTGCGCAAGGAGAGCGAGGCGAGCTTTTGGGAATCCAGGACTGTATCTCGAA 603
QY 829 CGTCTGTGATTAACCTCGACATTAATGAAGTGCAGTCTTGGCGATCACACTGGAGAA 888
Db |||
QY 604 AAATTTATCGATCCCGCCAGCGAGTGTGAATTCAGATCTTGGCCGATGCTTACGCGAAT 663
QY 889 GTTGTACACCTTTATGAACGTGACTGCTCACTGAGCGTCTGTCACCAAAAGTTGCGAA 948
Db |||
QY 664 GTAGTGATCTAGGAGGAGCGGATTTGCTCCATCAAGTCTGACCAAAAGTGTCTGAA 723
QY 949 ATTGCGGACGACAGATTTGGATTCAGAACTGGGTGATCGCATTTGTGCGGATGAGTA 1008
Db |||
QY 724 GAAGCCCCCAGTCCGCGCTATCGGCAGACCTGCGGAGAAATGGCGGATGCGCGCGTC 783
QY 1009 AAGTTCTGCGCTCCATTTGTTACAGGCGCGGAAACCGTGAATTTCTTGGTCGATGAA 1068
Db |||
QY 784 AAGTCTGTCAAGCGATCGGCTATCATCGTTCGCGCAGCGGTGAGTTCTGTGCGATGCG 843
QY 1069 AAGGGCAACAGCTTTCATCGAAATGAAACCCAGTATCCAGTTGAGCAGACCGTGACT 1128
Db |||
QY 844 ACCGGCAACTTCTACTTTCATGGAGATGAATACCCGCTATCCAAAGTCGAGCATCAGTACA 903
Db |||
```

```
QY 1129 GAAGAAGTCAACGAGGTGACCTGGTGAAGGCGGAGATGCGCTTGGCTGCTGTCACAC 1188
Db |||
QY 904 GAATGATTTACGGGACTGACTTGTGTTGGGAGCAGATTCGGATTGCCCAAGGCG--- 958
Db |||
QY 1189 TTGAAGGAATTTGGTCTGAGCCCAAGATAGATCAAGACCCAGGTGCAGCACTGAGTGC 1248
Db |||
QY 959 ----AAGCGTTCGCTTCCGCAAGCCGATATTCAACTGCGCGGCGATCGGATCGAATGC 1014
QY 1249 CGCATCACCAAGGAAGATCCAAACACGCGTTCGCGCCAGATACCGGAATATCACCGCG 1308
Db |||
QY 1015 CGTATCAATGCGGAAGATCGGATATCAATTTCCGCGCGAATCTTCCGCGCATTACAGCG 1074
QY 1309 TACCGCTACCAAGCGGAGTGGCTGCTTGAAGCGTGCAGTTCAGTTCGTTGCGGAA 1368
Db |||
QY 1075 TATTTACCGCGCGCGCGCGCTGCTGCTGATTCCTCATTTCCATGTTTATACCGACTACGAA 1134
QY 1369 ATCACCGCACATTTGACTCCATGCTGCTGATAAATGACCTGCGTGGTTCGACTTTGAA 1428
Db |||
QY 1135 ATTCCGCGCTTATACGATTCGCTGATGGCAATTTGATGTTGCGGTTGCAACCGGAA 1194
QY 1429 ACTGCTGTTGCTCGTGCACAGCGCGGTTGGCTGAGTTCACCGTGTCTGTTGTCACAC 1488
Db |||
QY 1195 GAGCGCATCGCGGAGTGCAGCGTGTCTGCGGGAATGCGCCATCACCGGCTTGGCGAG 1254
QY 1489 AACATGCTTTCTTGGTGGTGGTGGTGGGGAAGAGACTTCACTTCCAAAGCGCATCGCG 1548
Db |||
QY 1255 ACCTTAGTTTCCATGAGTGAATGTCAGATGCTGCTGCGGCGGAGTTCCTGCTGCGGGGAGTCTAT 1314
QY 1549 ACCGATTCATTG 1561
Db |||
QY 1315 ACCAACTTTGTTG 1327
Db |||
```

## RESULT 13

US-08-485-607-5

Sequence 5, Application US/08485607

Patent No. 5792627

GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki

TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA

TITLE OF INVENTION: Carboxylase

NUMBER OF SEQUENCES: 116

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White &amp; Durkee

STREET: 321 No. 5792627th Clark Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,607

FILING DATE: 07-JUN-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/956,700

FILING DATE: 10/21/92

ATTORNEY/AGENT INFORMATION:

NAME: Thomas E. No. 5792627thrup

REGISTRATION NUMBER: 33,268

REFERENCE/DOCKET NUMBER: ARCD:058

TELECOMMUNICATION INFORMATION:

TELEPHONE: 1-312-744-0090

TELEFAX: 1-312-755-4489

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1362 base pairs

TYPE: Nucleic acid



STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Oligonucleotide  
US-08-485-607-5

Query Match  
Best Local Similarity 53.8%; Score 313; DB 1; Length 1362;  
Matches 718; Conservative 0; Mismatches 600; Indels 15; Gaps 3;

1189 TTGAAGGAATGGGTCTGACCAAGATAGATCAAGACCAAGTCAAGACCAAGTCAAGACCAAGTCAAGTGC 1248  
959 ----AAGGCTGCGCTTCGGCAAGCCGATATTCACTGCGCGGCGCATCGCATCGAATGC 1014  
1249 CGCATCACCGAAGATCCAAACCGCTTCGCGCCAGATACCGAATCACTACCGG 1308  
1015 CGTATCAATGCGAAGATCCGGAATACAAATTTCCGCGCAATCCCTGGCGCATACAGC 1074  
1309 TACCGCTCACCAAGGCGAGCTGCGCTTCTGTGACGGTGCAGCTCAGCTCGGTGGCGAA 1368  
1075 TATTACCGCCCGCGCGCGCTTCTGTGCGATTCCCATGTTTATACCGACTACGAA 1134  
1369 ATCAGCGACATTTGACTCCATGCTGTGAAATGACTGCGGTGTTCCGACTTTGAA 1428  
1135 ATTCCGCGCTTATACGATTTCGCTGATTCGCAAAATGATTGTTCTGGGTGCAACACGGGAA 1194  
1429 ACTGCTGTGCTCGTGCAAGCGCGCTTGGTGTGAGTTTACCGTGTCTGTGGTGTGCAACC 1488  
1195 GAGGCGATCGCGCGATGCGAGTGTCTCTCGGGAATCGCCATCACCGGCTTGCAGC 1254  
1489 AACATTGGTTCTTGGTGTGCTGCTGCGGGAAGAGACTTCATTCAGCGCATCGCC 1548  
1255 ACCCTTAGTTTCCATCAGCTGATTTGCAGATGCTCTGCGGGAATCGCCGCGGAACTCTAT 1314  
1549 ACOGGAATCATG 1561  
1315 ACCAATTTGTTG 1327

RESULT 14  
US-08-475-879-5  
; Sequence 5, Application US/08475879  
; Patent No. 5972644  
; Patent No. 5972644 5786170  
; GENERAL INFORMATION:  
; APPLICANT: Robert Haselkorn and Piotr Gornicki  
; TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
; TITLE OF INVENTION: Carboxylase  
; NUMBER OF SEQUENCES: 116  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold White & Durkee  
; STREET: 321 No. 5972644 5786170th Clark Street  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60610  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII-DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/475,879  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/956,700  
; FILING DATE: 10/21/92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Thomas E. No. 5972644 5786170thrup  
; REGISTRATION NUMBER: 33,268  
; REFERENCE/DOCKET NUMBER: ARCD:058  
; TELEPHONE: 1-312-744-0090  
; TELEFAX: 1-312-755-4489  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1362 base pairs  
; TYPE: Nucleic acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Oligonucleotide  
US-08-475-879-5

Query Match 8.8%; Score 313; DB 2; Length 1362;  
Best Local Similarity 53.9%; Pred. No. 5,7e-79;  
Matches 718; Conservative 0; Mismatches 600; Indels 15; Gaps 3;

Qy 232 TTCAAAAGATCTTGGTAGCAACACCGCGCGAAATCGCGGTCCGTGCTTTCCGTGCGAGCA 291  
Db 7 TTCAACAGATCTTGATCGCAATCGCGGAAATCGCCCTGGCATCTCCGCACTTGT 66

Qy 292 CTCGAAACCGGTGCGACGACGAGTAGTATTTACCCCGGTGAAGATCGGGGATCATTCAC 351  
Db 67 CAAGAACTCGGGATCGGACGATCGCGGTTCATCCACTGTGGATCGCAACCGCTCCAT 126

Qy 352 CGCTCTTTTCTCTGAAGCTGTCCCAATGGTACCGAAGGCTCACCACTCAAGCGGTAC 411  
Db 127 GTGAGTACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 183

Qy 412 CTGACATCGATGAATTTATCGGTGACAGTAAAGATTAAGCAGATGCAATTTACCG 471  
Db 184 CTCAATATCCCAACATCATTTGCGCGGCGCTGACCCCTTAATGCGAGCGCAATTCACCC 243

Qy 472 GGATACGGCTTCTGCTGAAATGCCAGCTTGGCCGAGTGTGCGGAAACCGCAAT 531  
Db 244 GGTATGGCTTCTTGGCGGAGATGCGCTTTGAGAAATCTGCGCGATCACCATCTC 303

Qy 532 ACTTTTATGGCCCAACCCAGAGGTTCTGTATCTCACCGGTGATTAAGTCTCGGG---G 588  
Db 304 ACCTTTATGGCCCAACCCAGAGGTTCTGTATCTCACCGGTGATTAAGTCTCGGG---G 588

Qy 589 GTAACCGCCGCGAAGGCTGTGCTGCAAGTCTTTCGCGAATCCACCCGAGCAAAAC 648  
Db 364 GAAACATGACGAGGCTGCGGTTCGAGTTCGAGCATGCGGCGAGTACGGTCTGTGCGAT 423

Qy 649 ATGATGAGATCTTAAAGCGTGAAGCGACAGCTTACCCCATCTTTGTGAAGCGATT 708  
Db 424 GTTGATTCGCTGCAAGGTTGCTGCGAGATCGGCTATCCGCTCATGATCAAGCGAG 483

Qy 709 GCGGTGTTGCGGACGCGGTATGCTTTTGTGCTTCACTGATGAGTTCGCAATTA 768  
Db 484 GCGGGGCGGTGCTGCGGTATGCGGTGCTGCGGTGCTGCGGTGCTGCGGTGCTGCGGT 543

Qy 769 GCAACAGAAGATCTCTGTAAGTGAAGCGGCTTTCGCGATGCGCGGTATATGTCGAA 828  
Db 544 TTCCTTCTGCCAAGGAGAGCGAGGAGCTTTTGGGAATCCAGAGCTGTATCTCGAA 603

Qy 829 CGTCTGTGATTAACCTCAGATATGAGTGCAGATCCTTGGCGATCACACTGAGAA 888  
Db 604 AATTTATCGATCGCCCAACCGCTTGTGAATTTGAGATCTTGGCGATGCTTACGCAAT 663

Qy 889 GTTGTACACCTTTATGAACGTGACTGCTCACTGACGCGTCTCACCAAAAGTTGTCGAA 948  
Db 664 GTAGTGATCTAGGCGAGCGGATTTGCTCATTTCAAGCTGCTCACCAAAAGTGTGCGAA 723

Qy 949 ATTGCGGACGACAGCATTTGATCCAGACTCGGTGATCGCATTTGTGCGATGCA 1008  
Db 724 GAAGCCCCAGTCCGGCTGATCGGACAGCTCGCGGAGAAATGCGCGATGCGCGCTC 783

Qy 1009 AAGTTCTGCGCTCATTTGTTTACAGGCGCGGGAACCGTGAATTTCTTGTGCGATGAA 1068  
Db 784 AAGTCTCAAGCGATCGGCTCATCGGTGCGGACCGTGGAGTTCTTGTGCGATGCG 843

Qy 1069 AAGGGCAACAGCTTTCATCGAAATGAACCCAGTATCCAGTTGAGCAGACCGGTGAT 1128  
Db 844 ACCGGCAACTTCTACTTTCATGGAGATGAATACCGCATCCCAAGTCAGATCCAGTCA 903

Qy 1129 GAAGAGTCAACGAGTGGACCTTGTGAAGCGCAGATGCGTGTGCTGCTGTCGCAAC 1188  
Db 904 GAATGATACGGAGCTGGATTTGATGCGGAGAGATTCGATTCGCCAAGCG--- 958

Qy 1189 TTGAAGGAATTTGGTCTGACCCAGATGAAGATCAAGACCCAGCGTGCAGCATGCGAGTGC 1248  
Db 959 -----AAGCGTGTGCTTCCGCGAAGCGGATATTTCACTGCGCGGCGATGCGATGCG 1014

RESULT 15  
US-09-433-043B-5  
; Sequence 5, Application US/09433043B  
; Patent No. 6399342  
; GENERAL INFORMATION:  
; APPLICANT: HASELKORN, ROBERT  
; APPLICANT: GORNICKI, PIOTR  
; TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE  
; FILE REFERENCE: ARCD:3380US  
; CURRENT APPLICATION NUMBER: US/09/433,043B  
; PRIOR FILING DATE: 1999-10-25  
; PRIOR APPLICATION NUMBER: 08/475,879  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: 07/956,700  
; NUMBER OF SEQ ID NOS: 128  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1362  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-433-043B-5

Query Match 8.8%; Score 313; DB 4; Length 1362;  
Best Local Similarity 53.9%; Pred. No. 5,7e-79;  
Matches 718; Conservative 0; Mismatches 600; Indels 15; Gaps 3;

Qy 232 TTCAAAAGATCTTGGTAGCAACACCGCGCGAAATCGCGGTCCGTGCTTTCCGTGCGAGCA 291  
Db 7 TTCAACAGATCTTGATCGCAATCGCGGAAATCGCCCTGGCATCTCCGCACTTGT 66

Qy 292 CTCGAAACCGGTGCGACGACGAGTAGTATTTACCCCGGTGAAGATCGGGGATCATTCAC 351  
Db 67 CAAGAACTCGGGATCGGACGATCGCGGTTCACCTCACTGTGGATCGCAACCGCTCCAT 126

Qy 352 CGCTCTTTTCTGAGCTGTCCGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411  
Db 127 GTGAGTACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 183

Qy 412 CTGACATCGATGAATTTATCGGTGACAGTAAAGATTAAGCAGATGCAATTTACCG 471  
Db 184 CTCAATATCCCAACATCATTTGCGCGGCGCTGACCCCTTAATGCGAGCGCAATTCACCC 243

Qy 472 GGATACGGCTTCTGCTGAAATGCCAGCTTGGCCGAGTGTGCGGAAACCGCAAT 531  
Db 244 GGTATGGCTTCTTGGCGGAGATGCGCGCTTTGAGAAATCTGCGCGGCGATCACCATCTC 303

Search completed: March 23, 2004, 18:39:49  
Job time : 257 secs

```
QY 532 ACTTTATGCCCCAACCCACAGAGTTCTTGATCTCACCGGTGATAGTCTCCGCG---G 588
D 533 |||||
D 304 ACCTTTATGCCCCACGCCCGGATTCGATTTCGAGCCATGGGCGATAAATCCACCGCTAAG 363
QY 589 GTRACCCCGCGGAGAGAGGCTGGTCTCCAGTTTTCGCGGAATCCACCCCGAGCAAAAC 648
D 590 |||||
D 364 GAAACAATGACGGGTCGGGTTCCGACGATTCGGGCGAGTGACGGTCTGCTGACGGAT 423
QY 649 ATCGATGAGATCGTTAAAGCGCTGAAGCGCAGACTTACCCCATCTTTGTGAAGGCAGTT 708
D 650 |||||
D 424 GTTGATTGGCTGCCAAAGTTGTCGCGAGATCGGCTATCCCGTCATGATCAAAAGCGACG 483
QY 709 GCCGGTGGTGGCGAGCGGATGCGTTTGTGCTTCACTGATGAGCTTCGCAATTA 768
D 710 |||||
D 484 CGGGGGGGCGGTGTCGGGTATCGCGTATCGCGTGTGCGGTGACCTCGAGATCTGGAATACTG 543
QY 769 GCAACAGAAAGCATCTCGTGAAGCTGAAGCGCTTTTCGGCGATGGCGGCTATATGTCGAA 828
D 770 |||||
D 544 TTCCTTGTGCCCCAAGAGAGAGCGGAGAGCTTTTGGGAATCCAGGACTGTATCTCGAA 603
QY 829 CGTGTGTGANTAACCTTCAGCATTTGAAGTGCAGATCCTTGGCGATCACACTGGAGAA 888
D 830 |||||
D 604 AAATTTATCGATCGCCCCACGCCACGTTGAATTCAGATCTTGGCCGATGCTTACGGCAAT 663
QY 889 GTTGTAACCTTTATGAACGTGACTGCTCACTGACGCGTGTCTACCAAAAAGTTGCGAA 948
D 890 |||||
D 664 GTAGTGATCTAGCGGAGCGGATTTGCTCCANTTCAAGTGTCAACAAAGCTGCTCGAA 723
QY 949 ATTGGCCAGCAGATTTGGATCCAGAACTGCGTGATCGCATTTGCGGATGCGATTA 1008
D 950 |||||
D 724 GAAGCCCCCAGTCCGGCGCTATCGGCAGACCTGGCGCAGAAATGGGCGATGGCGCGTC 783
QY 1009 AAGTTCTCGCGCTCCATTGGTTTACCAGGGCGCGGAAACCGTGGAAATCTTGGTGCATGAA 1068
D 1010 |||||
D 784 AAAGTCGCTCAAGCGATCGGCTACATCGTCCGCGACCGTGGAGTTTCTGTTGATGCG 843
QY 1069 AAGGCAACACGCTTTCATCGAAATGAACCCAGTATCCAGGTTGAGCACACGCTGACT 1128
D 1070 |||||
D 844 ACCGGCAACTTCTACTTCATGAGATGAATACCGGCATCCAAATCGAGCATCCAGTCACA 903
QY 1129 GAAGAAGTCAAGGAGTGGACTGTTGAAGCGGAGATGGCTTGGTCTGTTGTCGAACC 1188
D 1130 |||||
D 904 GAATGATTAAGGAGTGGACTTGATTGCGGAGCAGATTCGGATTGCCCAAGGCG----- 958
QY 1189 TTGAAGGAATTTGGTCTGACCCCAAGATAAGATCAAGACCCACGGTGCAGCATGCACTGC 1248
D 1190 |||||
D 959 ----AAGCGCTGGCTTCGGGCAAGCGATATTCAACTGCGCGGCCATCGATCGAATGC 1014
QY 1249 CGCATCACCGAGAGATCCAAACACCGCTTCGCGCCAGATACGGNACTATCACCGCG 1308
D 1250 |||||
D 1015 CGTATCAATGCGGAAGATCCGGATACAAATTCGGCGCGAATCCTGGCGCGATTTACAGGC 1074
QY 1309 TACCGCTCACCGAGCGAGCTGGCGTTCGTTGACGGTGCAGCTCAGCTCGTGGCGGAA 1368
D 1310 |||||
D 1075 TATTTACCGCGCGCGCGCGGCTTCGTGCTGATTCCTCCATGTTTATACCGACTACGAA 1134
QY 1369 ATCAGCGCACATTTGATTCATGCTGTTGAATGACTGCTCGCGTGGTTCCGACTTTGAA 1428
D 1370 |||||
D 1135 ATTCGCGCCTATTACGATTCGCTGATTCGCAATTTGATTGTTGGGGTGCAACACCGGAA 1194
QY 1429 ACTGCTGTGCTGTCGACAGCGCGGTTGGCTGAGTTTCACTGCTGCTGGTGTTCGAACC 1488
D 1430 |||||
D 1195 GAGGCGATCGCGGATCGAGGCTGCTCTGCGGGAATCGGCCATCACCGGCTTCCGCGAG 1254
QY 1489 AACATTGGTTTCTTCGCTGCGGTGCTGCGGGAAGAGGACTTTCACCTCCAGCGCATCGCC 1548
D 1490 |||||
D 1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCTGAGTTTCTGCGCGGGGAATCTTAT 1314
QY 1549 ACCGATTTCATTG 1561
D 1550 |||||
D 1315 ACCAATTGTTG 1327
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 20:24:42 ; Search time 1213 Seconds  
(without alignments)  
11042.568 Million cell updates/sec

Title: US-10-045-072-1  
Perfect score: 3621  
Sequence: 1 tgggggggtagatcctg.....tgatgctgctgttcctaa 3621

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09C\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3621	100.0	3621	14	US-10-045-072-1
2	3621	100.0	3309400	9	US-09-738-626-1
3	3420	94.4	3420	9	US-09-738-626-1
4	3398.8	93.9	3474	9	US-09-974-973-1
5	3398.8	93.9	3474	9	US-09-974-973-3
6	1927.4	53.2	3423	12	US-10-282-122A-17701
7	1361.6	37.6	3381	12	US-10-282-122A-25766
8	1294.4	35.7	3384	12	US-10-282-122A-28626
9	1292.2	35.7	3381	12	US-10-282-122A-26500
10	681	18.8	3345	9	US-09-917-800A-1566
11	677.4	18.7	4017	9	US-09-880-107-3029
12	659	18.2	3432	15	US-10-369-493-31453
13	643.6	17.8	3459	15	US-10-369-493-35714
14	617	17.0	3441	15	US-10-369-493-35137
15	617	17.0	3453	15	US-10-369-493-38504

16	617	17.0	3453	15	US-10-369-493-38693	Sequence 38693, A
17	601.8	16.6	3453	15	US-10-369-493-41025	Sequence 41025, A
18	598	16.5	4152	15	US-10-398-221-3692	Sequence 3692, Ap
19	582	16.1	3414	15	US-10-369-493-42038	Sequence 42038, A
20	577.2	15.9	3441	12	US-10-282-122A-24304	Sequence 24304, A
21	572.8	15.8	3447	15	US-10-369-493-46799	Sequence 46799, A
22	565.2	15.6	3537	15	US-10-369-493-45626	Sequence 45626, A
23	565	15.6	1163020	15	US-10-398-221-10	Sequence 10, Appl
24	565	15.6	3011208	15	US-10-398-221-2058	Sequence 2058, Ap
25	549.4	15.2	3540	15	US-10-369-493-25178	Sequence 25178, A
26	548.8	15.2	3501	15	US-10-369-493-46506	Sequence 46506, A
27	548.8	15.2	3555	15	US-10-369-493-26175	Sequence 26175, A
28	547.2	15.1	2709	15	US-10-369-493-37987	Sequence 37987, A
29	522	14.4	3444	12	US-10-282-122A-10682	Sequence 10682, A
30	514	14.2	3429	9	US-09-815-242-6709	Sequence 6709, Ap
31	509.2	14.1	3426	12	US-10-282-122A-6344	Sequence 6344, Ap
32	500.4	13.8	20072	9	US-09-070-927A-89	Sequence 89, Appl
33	482.4	13.3	3426	12	US-10-282-122A-21758	Sequence 21758, A
34	479.8	13.3	5030	8	US-08-781-986A-324	Sequence 324, App
35	479.8	13.3	5030	12	US-10-329-624-324	Sequence 324, App
36	479.6	13.2	3453	12	US-10-282-122A-8207	Sequence 8207, Ap
37	478.2	13.2	3441	9	US-09-815-242-4413	Sequence 4413, Ap
38	462.6	12.8	3435	12	US-10-282-122A-15904	Sequence 15904, A
39	449.4	12.4	3432	12	US-10-282-122A-16593	Sequence 16593, A
40	421.8	11.6	3222	9	US-09-815-242-8263	Sequence 8263, Ap
41	420.6	11.6	3429	12	US-10-282-122A-16733	Sequence 16733, A
42	363.8	10.0	1362	15	US-10-369-493-36893	Sequence 36893, A
43	358.4	9.9	3230	15	US-10-369-493-37518	Sequence 37518, A
44	337.2	9.3	4161	15	US-10-369-493-27492	Sequence 27492, A
45	311.4	8.6	1362	9	US-09-767-479-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-10-045-072-1  
; Sequence 1, Application US/10045072  
; Publication No. US20030027305A1  
; GENERAL INFORMATION:  
; APPLICANT: Sinskey, Anthony J.  
; APPLICANT: Lessard, Philip A.  
; APPLICANT: Willis, Laura B.  
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum  
; FILE REFERENCE: 1533.0790002  
; CURRENT APPLICATION NUMBER: US/10/045,072  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: US 09/677,575  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: US 09/220,081  
; PRIOR FILING DATE: 1998-12-23  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3621  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (199)...(3621)  
US-10-045-072-1

Query Match	100.0%;	Score 3621;	DB 14;	Length 3621;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3621;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	TCGGGGGGGGTGTAGATCCCTGGGGGGTATTATTCATTCACTTTGGCTTGAAGTCGTCAGG	60	
Db	1	TCGGGGGGGGTGTAGATCCCTGGGGGGTATTATTCATTCACTTTGGCTTGAAGTCGTCAGG	60	
Qy	61	TCAGGGAGGTGTGCGGAAAAACATTGAGAGGAAAAACAAAAACCGATGTTGATTCGGGG	120	
Db	61	TCAGGGAGGTGTGCGGAAAAACATTGAGAGGAAAAACAAAAACCGATGTTGATTCGGGG	120	

QY	121	AATCGGGGTTACGATACTAGGACGCAGTGA	CTGCTATCACCTTTGGCGGTCTCTGTTG	180
DB	121	AATCGGGGTTACGATACTAGGACGCAGTGA	CTGCTATCACCTTTGGCGGTCTCTGTTG	180
QY	181	AAAGGATANTTACTCTAGTGTGCACTCACAT	CTTCAAGCTTCCAGCATTTCAAAAG	240
DB	181	AAAGGATANTTACTCTAGTGTGCACTCACAT	CTTCAAGCTTCCAGCATTTCAAAAG	240
QY	241	ATCTTGGTAGCAAAACCGCGCGAAATTCGG	GTCTTTCCGTGACGACTCGAAACC	300
DB	241	ATCTTGGTAGCAAAACCGCGCGAAATTCGG	GTCTTTCCGTGACGACTCGAAACC	300
QY	301	GGTGACGACGGTAGCTATTTTACCCCGGT	GAAGATCGGGATCATTTCCACCGCTCTTTT	360
DB	301	GGTGACGACGGTAGCTATTTTACCCCGGT	GAAGATCGGGATCATTTCCACCGCTCTTTT	360
QY	361	GCTTCTGAAGCTGTCCGCAATGGTAGCGA	GGTCCAGTCCCGGATCATTTACCGGATCGGC	420
DB	361	GCTTCTGAAGCTGTCCGCAATGGTAGCGA	GGTCCAGTCCCGGATCATTTACCGGATCGGC	420
QY	421	GATGAAATTAATCGGTGCAGCTAAAAAGT	TAAAGCAGATGCCATTTACCGGGATACGGC	480
DB	421	GATGAAATTAATCGGTGCAGCTAAAAAGT	TAAAGCAGATGCCATTTACCGGGATACGGC	480
QY	481	TTCTGTCTGAAATATGCCAGCTTCCGCGA	GTGTGCGGAAAAACGGCATTACTTTTAT	540
DB	481	TTCTGTCTGAAATATGCCAGCTTCCGCGA	GTGTGCGGAAAAACGGCATTACTTTTAT	540
QY	541	GGCCCAACCCACAGAGTCTTTTGATCTCA	CCGGTGATAAGTCTCCGCGGTAAACCGCGCG	600
DB	541	GGCCCAACCCACAGAGTCTTTTGATCTCA	CCGGTGATAAGTCTCCGCGGTAAACCGCGCG	600
QY	601	AAGAAGCTGTCTGCAGTTTGTGCGGATTC	ACCCATCTTTGTGAAGCAGTTGCGGTGTGGC	660
DB	601	AAGAAGCTGTCTGCAGTTTGTGCGGATTC	ACCCATCTTTGTGAAGCAGTTGCGGTGTGGC	660
QY	661	GTAAAAAGCGCTGAAGCCAGACTTAACCC	ATCTTTGTGAAGCAGTTGCGGTGTGGC	720
DB	661	GTAAAAAGCGCTGAAGCCAGACTTAACCC	ATCTTTGTGAAGCAGTTGCGGTGTGGC	720
QY	721	GGACGCGGTATGCGTTTGTGTTTCACTG	ATGAGCTTCGCAATTAGCAACAGAAC	780
DB	721	GGACGCGGTATGCGTTTGTGTTTCACTG	ATGAGCTTCGCAATTAGCAACAGAAC	780
QY	781	TCTCGTGAAGCTGAAGCGGCTTTCCGCG	ATGCGCGGTATATGTCGAACTGTGTGAT	840
DB	781	TCTCGTGAAGCTGAAGCGGCTTTCCGCG	ATGCGCGGTATATGTCGAACTGTGTGAT	840
QY	841	AACCTCAGCATATTTGAAGTCAGATCC	TTTGGGATCACACTGGAGAGTTGACACCT	900
DB	841	AACCTCAGCATATTTGAAGTCAGATCC	TTTGGGATCACACTGGAGAGTTGACACCT	900
QY	901	TATGAACTGACTGCTCACTGACGCTCG	TACCAAAAAGTTGCGAAATTCGCCGAC	960
DB	901	TATGAACTGACTGCTCACTGACGCTCG	TACCAAAAAGTTGCGAAATTCGCCGAC	960
QY	961	CAGCATTTGGATCCAGACTCGGTGATCG	CAATTTGTGCGGATGCAAGTTCTGCCCG	1020
DB	961	CAGCATTTGGATCCAGACTCGGTGATCG	CAATTTGTGCGGATGCAAGTTCTGCCCG	1020
QY	1021	TCCATTGGTTACCGGGCGCGGAAACCG	TGGATTTCTTGTGCGATGAAAGGCAACC	1080
DB	1021	TCCATTGGTTACCGGGCGCGGAAACCG	TGGATTTCTTGTGCGATGAAAGGCAACC	1080
QY	1081	GTCCTCATCGAAATGAAACCCAGTATCC	AGGTTGAGCACACCGTGAAGAGTCAACC	1140
DB	1081	GTCCTCATCGAAATGAAACCCAGTATCC	AGGTTGAGCACACCGTGAAGAGTCAACC	1140
QY	1141	GAGGTGGAACCTGGTGAAGGCGCAGATG	CGCTTGGCTGTGGTGCAACTTTGAAGGA	1200
DB	1141	GAGGTGGAACCTGGTGAAGGCGCAGATG	CGCTTGGCTGTGGTGCAACTTTGAAGGA	1200

Qy	1201	GGTCTGACCCAGATAAGATCAAGACCCACGGTGCAGCTCCGATGCGCGATCAACG	1260
Db	1201	GGTCTGACCCAGATAAGATCAAGACCCACGGTGCAGCTCCGATGCGCGATCAACG	1260
Qy	1261	GAAGATCCAAACAAACGGCTTCGGCCAGATACCGAACTATCAACCGGTACCGCTCAACA	1320
Db	1261	GAAGATCCAAACAAACGGCTTCGGCCAGATACCGAACTATCAACCGGTACCGCTCAACA	1320
Qy	1321	GGCGGAGCTGGCGTTGGTCTTGACGGTGCAGCTCAGCTCGGTGGGGAATCAACCGCACAC	1380
Db	1321	GGCGGAGCTGGCGTTGGTCTTGACGGTGCAGCTCAGCTCGGTGGGGAATCAACCGCACAC	1380
Qy	1381	TTTGACTCCATGCTGGTGAATAATGACCTCGCGTGGTTCGGACTTTGAAACTGCTGTGCT	1440
Db	1381	TTTGACTCCATGCTGGTGAATAATGACCTCGCGTGGTTCGGACTTTGAAACTGCTGTGCT	1440
Qy	1441	CGTGCAACGCGCGCTTGGCTGAGTTTACCGTGTCTGGTGTGTAACCAACAACTGGTGTTC	1500
Db	1441	CGTGCAACGCGCGCTTGGCTGAGTTTACCGTGTCTGGTGTGTAACCAACAACTGGTGTTC	1500
Qy	1501	TTGCGTGGCTGTGCGGGAAGAGGACTTCACTTCCAAAGCGGATCGCCACCGATTTCAAT	1560
Db	1501	TTGCGTGGCTGTGCGGGAAGAGGACTTCACTTCCAAAGCGGATCGCCACCGGATTTCAAT	1560
Qy	1561	GCGGATCAACCGCACCTTCTTCAAGGCTCCACCTGCTGATGATGAGCAGGAGCATCTCTG	1620
Db	1561	GCGGATCAACCGCACCTTCTTCAAGGCTCCACCTGCTGATGATGAGCAGGAGCATCTCTG	1620
Qy	1621	GATTACTTGCGAGATGTACCGGTGAACAAGCTCTATGTTGGCTTCCAAAGGATGTTGCA	1680
Db	1621	GATTACTTGCGAGATGTACCGGTGAACAAGCTCTATGTTGGCTTCCAAAGGATGTTGCA	1680
Qy	1681	GCTCCTATCGATAAGCTGCTTAACATCAAGGATCTGCCACTGCGCAGCGGTTCCCGTGAC	1740
Db	1681	GCTCCTATCGATAAGCTGCTTAACATCAAGGATCTGCCACTGCGCAGCGGTTCCCGTGAC	1740
Qy	1741	GCGCTGAAGCAGCTTGGCCACGCGCGTTTGGCTCTGATCTCCGTGAGCAGGACGCACTG	1800
Db	1741	GCGCTGAAGCAGCTTGGCCACGCGCGTTTGGCTCTGATCTCCGTGAGCAGGACGCACTG	1800
Qy	1801	GCAGTTACTGATACCACTTCCGCGATGCAACAGTCTTTGCTTGGACCCCGAGTCCGC	1860
Db	1801	GCAGTTACTGATACCACTTCCGCGATGCAACAGTCTTTGCTTGGACCCCGAGTCCGC	1860
Qy	1861	TCATTCGCACTGAAGCCTTGGCGGAGAGCGCGTGCAGAGCTGACTCTGAGCTTTTGTCC	1920
Db	1861	TCATTCGCACTGAAGCCTTGGCGGAGAGCGCGTGCAGAGCTGACTCTGAGCTTTTGTCC	1920
Qy	1921	GTGAGAGCTTGGGCGCGGGAACCTACGATGTGGCGATGCGTTTCTCTTTGAGGATCCG	1980
Db	1921	GTGAGAGCTTGGGCGCGGGAACCTACGATGTGGCGATGCGTTTCTCTTTGAGGATCCG	1980
Qy	1981	TGGGACAGCTGCGAGAGCTTGGCGGAGCGGATGCGGATGTAAACATTCAGATGCTGCTT	2040
Db	1981	TGGGACAGCTGCGAGAGCTTGGCGGAGCGGATGCGGATGTAAACATTCAGATGCTGCTT	2040
Qy	2041	CGCGCGCGCAACACCGTGGGATACACCCCGTACCGACTCCGCTCGCGCGGCTTGGTT	2100
Db	2041	CGCGCGCGCAACACCGTGGGATACACCCCGTACCGACTCCGCTCGCGCGGCTTGGTT	2100
Qy	2101	AAGGAGCTGCGAGCTTGGCGTGGATCTTCCGATCTTCCGAGCGGCTTAAACAGCTC	2160
Db	2101	AAGGAGCTGCGAGCTTGGCGTGGATCTTCCGATCTTCCGAGCGGCTTAAACAGCTC	2160
Qy	2161	TCCAGATGCGTCCAGCAATCGACGAGTCTTGAGACCAACACCGCGGTAGCGAGGTG	2220
Db	2161	TCCAGATGCGTCCAGCAATCGACGAGTCTTGAGACCAACACCGCGGTAGCGAGGTG	2220
Qy	2221	GCTATGGCTTATCTGGTGATCTCTCTGATCCAAATGAAAGCTCTACACCTCGATTAC	2280
Db	2221	GCTATGGCTTATCTGGTGATCTCTCTGATCCAAATGAAAGCTCTACACCTCGATTAC	2280
Qy	2281	TACCTTAAAGATGGCAGAGGAGATCGTCAAGTCTCGGCTCTCATCTTTGGGCAATTAAGAT	2340

Db 2281 TACCTAAGATGCGAGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAAGAT 2340  
Qy 2341 ATGGCTGGTCTGCTTCCGCCAGTGGGTAACCAAGCTGGTCAACGCACTGGCGCGTGA 2400  
Db 2341 ATGGCTGGTCTGCTTCCGCCAGTGGGTAACCAAGCTGGTCAACGCACTGGCGCGTGA 2400  
Qy 2401 TTCCGATGCGCATGTCAGTGCACACCCAGCACTGGCGGGTGGCCAGTGGCAACCTAC 2460  
Db 2401 TTCCGATGCGCATGTCAGTGCACACCCAGCACTGGCGGGTGGCCAGTGGCAACCTAC 2460  
Qy 2461 TTTCGCTGCACTCAAGCTGGTGCAGATGCTGTGACGGTGTTCGCGCACCACTGTCTGGC 2520  
Db 2461 TTTCGCTGCACTCAAGCTGGTGCAGATGCTGTGACGGTGTTCGCGCACCACTGTCTGGC 2520  
Qy 2521 ACCACTCCAGCAATCCCTGTCTGCCATTTGTTGCTGCAATTCGCGCACACCCGTGCGAT 2580  
Db 2521 ACCACTCCAGCAATCCCTGTCTGCCATTTGTTGCTGCAATTCGCGCACACCCGTGCGAT 2580  
Qy 2581 ACCGGTTTGAAGCTCGAGGCTGTTTCTGACCTCGAGCGGTACTGGAGCAGTGGCGGA 2640  
Db 2581 ACCGGTTTGAAGCTCGAGGCTGTTTCTGACCTCGAGCGGTACTGGAGCAGTGGCGGA 2640  
Qy 2641 CTGTACCTGCCATTTAGTCTGAACCCAGGCCAACCCGGTCCGCTCTACCGCCACGAA 2700  
Db 2641 CTGTACCTGCCATTTAGTCTGAACCCAGGCCAACCCGGTCCGCTCTACCGCCACGAA 2700  
Qy 2701 ATCCAGCGGACAGTGTTCACCTCGTGCAGCCAGCCACCCAGCTGGGCCCTTGGGAT 2760  
Db 2701 ATCCAGCGGACAGTGTTCACCTCGTGCAGCCAGCCACCCAGCTGGGCCCTTGGGAT 2760  
Qy 2761 CGTTTCCGAATCATCGAAGCAACTACGACCGGTAAATGAGATGCTGGGACGCCCAACC 2820  
Db 2761 CGTTTCCGAATCATCGAAGCAACTACGACCGGTAAATGAGATGCTGGGACGCCCAACC 2820  
Qy 2821 AAGTCAACCCATCTCCAGGTTTGGGACCTCGCACTCCCACTCGTTGGTGGGGT 2880  
Db 2821 AAGTCAACCCATCTCCAGGTTTGGGACCTCGCACTCCCACTCGTTGGTGGGGT 2880  
Qy 2881 GTGATCCAGCAGACTTTGCTGCGCATCCCAAAAGTACGACATCCAGACTCTGTATC 2940  
Db 2881 GTGATCCAGCAGACTTTGCTGCGCATCCCAAAAGTACGACATCCAGACTCTGTATC 2940  
Qy 2941 CGTTTCTGCGCGGAGCTTGGTAAACCTCAGGTGGTGGCCAGCAGCACTGGCCACC 3000  
Db 2941 CGTTTCTGCGCGGAGCTTGGTAAACCTCAGGTGGTGGCCAGCAGCACTGGCCACC 3000  
Qy 3001 CGCGCACTGGAAGCGCTCCGAAGGCAAGGCACTCTGACGGAAGTTCTGAGGAAGAG 3060  
Db 3001 CGCGCACTGGAAGCGCTCCGAAGGCAAGGCACTCTGACGGAAGTTCTGAGGAAGAG 3060  
Qy 3061 CAGGCGCACTCGAGCTGATGATTTCCAGGAAGTTCGCAATAGCTCTCAACCGCTCTG 3120  
Db 3061 CAGGCGCACTCGAGCTGATGATTTCCAGGAAGTTCGCAATAGCTCTCAACCGCTCTG 3120  
Qy 3121 TTCCGGAAGCAACCGAAGAGTTCTCGAGCAGCCGTGGCCGCTTCCGCAACCTCTGG 3180  
Db 3121 TTCCGGAAGCAACCGAAGAGTTCTCGAGCAGCCGTGGCCGCTTCCGCAACCTCTGG 3180  
Qy 3181 CTGGATGATCGTGAATTTCTTACGGCTGGTTCGAAGCGCGCGAGACTTTGATCCGCTG 3240  
Db 3181 CTGGATGATCGTGAATTTCTTACGGCTGGTTCGAAGCGCGCGAGACTTTGATCCGCTG 3240  
Qy 3241 CCAGATGCGCACCCCACTGCTTCTGCTGATGATGATCTCTGAGCCAGAGATGAAG 3300  
Db 3241 CCAGATGCGCACCCCACTGCTTCTGCTGATGATGATCTCTGAGCCAGAGATGAAG 3300  
Qy 3301 GGTATGCGCAATTTGTGGCCAACTCAACGCGCAGATCCGCGCAATCGGTGGGTGAC 3360  
Db 3301 GGTATGCGCAATTTGTGGCCAACTCAACGCGCAGATCCGCGCAATCGGTGGGTGAC 3360  
Qy 3361 CGCTCGTGTGATCTGTGACCGCAACCCGAGAAAGGAGATTCCTCCCAAGGCGCAT 3420

Db 3361 CGCTCGTGTGATCTGTCAACCGCAACCGCAAGAAAGCAGATTCCTCCAAAGGCGCAT 3420  
Qy 3421 GTTGTCTCACATTCGCTGTGTGTGTCAACCGTCACTGTTGCTGAAGTGTATGAGTCAAG 3480  
Db 3421 GTTGTCTCACATTCGCTGTGTGTGTCAACCGTCACTGTTGCTGAAGTGTATGAGTCAAG 3480  
Qy 3481 GCTGAGATGCGATGCGCAATCATCGAGGCTATGAAGATGAAGCAACAATCACTGCTTCT 3540  
Db 3481 GCTGAGATGCGATGCGCAATCATCGAGGCTATGAAGATGAAGCAACAATCACTGCTTCT 3540  
Qy 3541 GTTACGCGCAAAATCGATCGCTGTGTGTCTGCTGCAAGAGTGGAGTGGCGAC 3600  
Db 3541 GTTACGCGCAAAATCGATCGCTGTGTGTCTGCTGCAAGAGTGGAGTGGCGAC 3600  
Qy 3601 TTGATCGTGTGCTGTTCTTAA 3621  
Db 3601 TTGATCGTGTGCTGTTCTTAA 3621

RESULT 2  
US-09-738-626-1  
; Sequence 1, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENO, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 1  
; LENGTH: 3309400  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-1

Query Match 100.0%; Score 3621; DB 9; Length 3309400;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TGGGGCGGGGTTAGATCTCTGGGGGTTATTTCAATTCATTTGGCTTGAAGTCTGTCAGG 60  
Db 705013 TGGGGCGGGGTTAGATCTCTGGGGGTTATTTCAATTCATTTGGCTTGAAGTCTGTCAGG 705072  
Qy 61 TCAGGGAGTGTTCGCCGAACATTTGAGAGAAACAAACCCGATGTTTCATTTGGGGG 120  
Db 705073 TCAGGGAGTGTTCGCCGAACATTTGAGAGAAACAAACCCGATGTTTCATTTGGGGG 705132  
Qy 121 AATCGGGGTTACGATACCTAGGACGAGTACTGCTATCACCTTGGCGGTCTCTTTGTTG 180  
Db 705133 AATCGGGGTTACGATACCTAGGACGAGTACTGCTATCACCTTGGCGGTCTCTTTGTTG 705192  
Qy 181 AAAGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240  
Db 705193 AAAGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 705252  
Qy 241 ATCTTGTGTAACAAACCGCGCGGAAATCGGGTCCGTGCTTTCCGTGACGACTCGAAACC 300





QY 2461 TTGCTGAGCTCAAGCTGGTGGAGATGCTGTTGACGGTGTCTTCGCGACCACTGTCGGC 2520  
DB 707473 TTGCTGAGCTCAAGCTGGTGGAGATGCTGTTGACGGTGTCTTCGCGACCACTGTCGGC 707532  
QY 2521 ACCACCTCCAGCCATCCCTGTCTGCGATGTTGCTGATTCGGGACACACCCGTGCGCAT 2580  
DB 707533 ACCACCTCCAGCCATCCCTGTCTGCGATGTTGCTGATTCGGGACACACCCGTGCGCAT 707592  
QY 2581 ACCGTTTGGAGCTGGAGCTGTTCTGACCTCCAGCCGCTACTGGGAAGCAGTGGCGGA 2640  
DB 707593 ACCGTTTGGAGCTGGAGCTGTTCTGACCTCCAGCCGCTACTGGGAAGCAGTGGCGGA 707652  
QY 2641 CTGTACTGCCATTTGATGCTGGAAACCCAGGCCCAACCCGCTGCGCTTACCGCCAGAA 2700  
DB 707653 CTGTACTGCCATTTGATGCTGGAAACCCAGGCCCAACCCGCTGCGCTTACCGCCAGAA 707712  
QY 2701 ATCCAGGCGGACAGTTGCTCAACCTGGTGCACAGGCCACCGCACTGGGCGCTTGGCGAT 2760  
DB 707713 ATCCAGGCGGACAGTTGCTCAACCTGGTGCACAGGCCACCGCACTGGGCGCTTGGCGAT 707772  
QY 2761 CGTTTCGAACCTCATCGAAGACAACTACGACCGCTTAATGAGATGCTGGGACGCCCAACC 2820  
DB 707773 CGTTTCGAACCTCATCGAAGACAACTACGACCGCTTAATGAGATGCTGGGACGCCCAACC 707832  
QY 2821 AAGTCAACCCATCTCTCAAGTTGTTGGCGACCTCGCACTCCACCTGCTGGTGGCGGT 2880  
DB 707833 AAGTCAACCCATCTCTCAAGTTGTTGGCGACCTCGCACTCCACCTGCTGGTGGCGGT 707892  
QY 2881 GTGATCCAGCAGACTTTGCTGCGGATCCACAAAGTACGACATCCAGACTCTGTGATC 2940  
DB 707893 GTGATCCAGCAGACTTTGCTGCGGATCCACAAAGTACGACATCCAGACTCTGTGATC 707952  
QY 2941 GGTTCCTGGGGGAGCTGTTGTAACCTCAGGTGGCTGGCCAGAGCCACTGGCGACC 3000  
DB 707953 GGTTCCTGGGGGAGCTGTTGTAACCTCAGGTGGCTGGCCAGAGCCACTGGCGACC 708012  
QY 3001 CGCGCACTGGAAGGCGCTCCGAAGGCAAGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060  
DB 708013 CGCGCACTGGAAGGCGCTCCGAAGGCAAGCACCTCTGACGGAAGTTCCTGAGGAAGAG 708072  
QY 3061 CAGGCGCACTGAGCTGATGATTCGAAGGAAGTGTGCAATAGCTCAACCGGCTGCTG 3120  
DB 708073 CAGGCGCACTGAGCTGATGATTCGAAGGAAGTGTGCAATAGCTCAACCGGCTGCTG 708132  
QY 3121 TTCCGGAAGCCAAACGGAAGAGTTCTCGAGCACCGTGGCGCTTCGGCAACACCTCTGCG 3180  
DB 708133 TTCCGGAAGCCAAACGGAAGAGTTCTCGAGCACCGTGGCGCTTCGGCAACACCTCTGCG 708192  
QY 3181 CTGGATGATCGTGAATCTTCTAGCGCTGTGCGAAGCCGCGAGACTTTGATCCGCTG 3240  
DB 708193 CTGGATGATCGTGAATCTTCTAGCGCTGTGCGAAGCCGCGAGACTTTGATCCGCTG 708252  
QY 3241 CCAGATGCGGACCCCACTGCTGTTGCTGGATGGATCTCTGAGCCAGAGATTAAG 3300  
DB 708253 CCAGATGCGGACCCCACTGCTGTTGCTGGATGGATCTCTGAGCCAGAGATTAAG 708312  
QY 3301 GGTATGCGCAATGTTGTGGCCAACTCAACGCGCAGATCCGCCCAATGCGTGGTGCAG 3360  
DB 708313 GGTATGCGCAATGTTGTGGCCAACTCAACGCGCAGATCCGCCCAATGCGTGGTGCAG 708372  
QY 3361 CGCTCGGTGATGCTGTCACCGCAACCGCAGAAAGGAGGATCTCTCCAAACAGGGCCAT 3420  
DB 708373 CGCTCGGTGATGCTGTCACCGCAACCGCAGAAAGGAGGATCTCTCCAAACAGGGCCAT 708432  
QY 3421 GTTGTGCAACCATTCGCTGGTGTGCTCACCCTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480  
DB 708433 GTTGTGCAACCATTCGCTGGTGTGCTCACCCTGACTGTTGCTGAAGGTGATGAGGTCAAG 708492  
QY 3481 GCTGAGATGCACTGCGCAATCATCGAGGCTATGAGATGGAAGCAACATCACTGCTTCT 3540  
DB 708493 GCTGAGATGCACTGCGCAATCATCGAGGCTATGAGATGGAAGCAACATCACTGCTTCT 708552

QY 3541 GTTGACGGCAAAATCGATFCGGTGTGGTTCCTGCTCAACGAAGGTGGAAGGTGGCGAC 3600  
DB 708553 GTTGACGGCAAAATCGATFCGGTGTGGTTCCTGCTCAACGAAGGTGGAAGGTGGCGAC 708612  
QY 3601 TTGATCGTGGTTCCTTCTTAA 3621  
DB 708613 TTGATCGTGGTTCCTTCTTAA 708633

RESULT 3  
US-09-738-626-765  
; Sequence 765, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: Patent in ver. 3.0  
; SEQ ID NO 765  
; LENGTH: 3420  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-765

Query Match 94.4%; Score 3420; DB 9; Length 3420;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 GTGTGCACTCACACATCTTCAACGCTTCCAGCATTCAGAAAGATCTTGTAGCAACCGC 258  
DB 1 GTGTGCACTCACACATCTTCAACGCTTCCAGCATTCAGAAAGATCTTGTAGCAACCGC 60  
QY 259 GCGAAATCGCGTCCGCTTTCGGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 318  
DB 61 GCGAAATCGCGTCCGCTTTCGGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 120  
QY 319 ATTTACCCCGGTGAGATCGGGATCATTCACCGCTCTTTTGGCTTCTGAAGCTGTCCGC 378  
DB 121 ATTTACCCCGGTGAGATCGGGATCATTCACCGCTCTTTTGGCTTCTGAAGCTGTCCGC 180  
QY 379 ATTGGTACCGAAGGTCCACGCTCAAGGGTACCTGGACATCGATGAATATTCGGTGCA 438  
DB 181 ATTGGTACCGAAGGTCCACGCTCAAGGGTACCTGGACATCGATGAATATTCGGTGCA 240  
QY 439 GCTAAAGGATTAAGCAGATGCCATTTACCCGGGATACGGCTTCTGCTGAATATGCG 498  
DB 241 GCTAAAGGATTAAGCAGATGCCATTTACCCGGGATACGGCTTCTGCTGAATATGCG 300  
QY 499 CAGCTTCCCGCGAGTGTGCGGAAACCGCATTAATTTATTTGGCCCAACCCGAGGTT 558  
DB 301 CAGCTTCCCGCGAGTGTGCGGAAACCGCATTAATTTATTTGGCCCAACCCGAGGTT 360  
QY 559 CTTGATCTACCGGTGATAGTCTCGCGGTAAACCGCCGGAAGAGCTGTCTGCCA 618  
DB 361 CTTGATCTACCGGTGATAGTCTCGCGGTAAACCGCCGGAAGAGCTGTCTGCCA 420

619 GTTTTGGCGGAATCCACCCCGAGCGAAACATCGATGATGATCGTTAAAGCGCTGAAGGC 678  
621 GTTTTGGCGGAATCCACCCCGAGCGAAACATCGATGATGATCGTTAAAGCGCTGAAGGC 480  
679 CAGACTTACCCCATCTTTGTGAAGGAGTTGCGCGTGTGGCGGACGCGGTATGCGTTT 738  
481 CAGACTTACCCCATCTTTGTGAAGGAGTTGCGCGTGTGGCGGACGCGGTATGCGTTT 540  
739 GTTGTTCACCTGATGAGCTTCGCAATTTAGCAACAGAGCATCTCGTGAAGCTGAAGCG 798  
541 GTTGTTCACCTGATGAGCTTCGCAATTTAGCAACAGAGCATCTCGTGAAGCTGAAGCG 600  
799 GCTTTTCGGCGATGGCGGTATATGTCGAACGTCGTGTGATTAACCCCTCAGCATATTGAA 858  
601 GCTTTTCGGCGATGGCGGTATATGTCGAACGTCGTGTGATTAACCCCTCAGCATATTGAA 660  
859 GTGCGATCTTGGCGATCACACTGAGAGAGTTGTACACCTTTATGAAAGTGTGATCTCA 918  
661 GTGCGATCTTGGCGATCACACTGAGAGAGTTGTACACCTTTATGAAAGTGTGATCTCA 720  
919 GTGCGAGTCGTCAACCAAAAGTTGTGAAATTTGGCCAGCACAGCATTTGGATCCAGAA 978  
721 GTGCGAGTCGTCAACCAAAAGTTGTGAAATTTGGCCAGCACAGCATTTGGATCCAGAA 780  
979 CTGCGTGAATCGCATTTGTGCGGATGAGTAAAGTTCTGCGGCTCCATTTGGTTACAGGGC 1038  
781 CTGCGTGAATCGCATTTGTGCGGATGAGTAAAGTTCTGCGGCTCCATTTGGTTACAGGGC 840  
1039 GCGGGAACCGTGGAAATCTTGTGATGAAAGGCGAACACAGTCTTCATCGAAATGAAC 1098  
841 GCGGGAACCGTGGAAATCTTGTGATGAAAGGCGAACACAGTCTTCATCGAAATGAAC 900  
1099 CCAAGTATCCAGGTTGAGCACACCGTGAAGTGTGAGAGTCAACCGAGTGGACCTGGTGAAG 1158  
901 CCAAGTATCCAGGTTGAGCACACCGTGAAGTGTGAGAGTCAACCGAGTGGACCTGGTGAAG 960  
1159 GCGGAGATCGCTGCTGCTGTCGCAACTTGTGAGAGTGTGCTGACCCAGATGAAG 1218  
961 GCGGAGATCGCTGCTGCTGTCGCAACTTGTGAGAGTGTGCTGACCCAGATGAAG 1020  
1219 ATCAAGACCCACCGTGCAGACTGCGCATCACCGAAGATCCAAACAGCGC 1278  
1021 ATCAAGACCCACCGTGCAGACTGCGCATCACCGAAGATCCAAACAGCGC 1080  
1279 TTCCGCGCAGATACCGGAATATCACCGGTACCGCTCACAGGCGGAGCTGGCTGCT 1338  
1081 TTCCGCGCAGATACCGGAATATCACCGGTACCGCTCACAGGCGGAGCTGGCTGCT 1140  
1339 CTTTGAAGGTCAGCTCAGCTCGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG 1398  
1141 CTTTGAAGGTCAGCTCAGCTCGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG 1200  
1399 AAAATGACCTGCGGTGCTCGACTTTGAAACTGTGCTGTGCTGCTGCAAGCGCGGTG 1458  
1201 AAAATGACCTGCGGTGCTCGACTTTGAAACTGTGCTGTGCTGCTGCAAGCGCGGTG 1260  
1459 GCTGAGTTACCGGTGCTGCTGTTGCAACCAATTTGTTTCTTGGCTGCTGCTGCTGCGG 1518  
1261 GCTGAGTTACCGGTGCTGCTGTTGCAACCAATTTGTTTCTTGGCTGCTGCTGCTGCGG 1320  
1519 GAAAGAGATTTCATCTTCAAGCGCATTCGCCACCGGATTCATTTGCCGATCACCGCACCTC 1578  
1321 GAAAGAGATTTCATCTTCAAGCGCATTCGCCACCGGATTCATTTGCCGATCACCGCACCTC 1380  
1579 CTTTCAAGGCTCCACTGCTGATGATGAGGAGGAGCATCTTGGATTTACTTGGCAGATGTC 1638  
1381 CTTTCAAGGCTCCACTGCTGATGATGAGGAGGAGCATCTTGGATTTACTTGGCAGATGTC 1440  
1639 ACCGTGAACAGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1698  
1441 ACCGTGAACAGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500

1699 CTTAAACATCAAGGATCTGCCACTGCGACGCGTTCCCGTGACCGCTGAAGCAGCTTGGC 1758  
1501 CTTAAACATCAAGGATCTGCCACTGCGACGCGGTTCCCGTGACCGCTGAAGCAGCTTGGC 1560  
1759 CCAGCCGCGTGTGCTGCTGATCTCCGTGAGCAGGACGCACTGCGAGTTACTGATACCAAC 1818  
1561 CCAGCCGCGTGTGCTGCTGATCTCCGTGAGCAGGACGCACTGCGAGTTACTGATACCAAC 1620  
1819 TTCCGCGATGCAACCAAGCTTTGCTTGGACCCGAGTCCGCTCATTTCCGACTGAAGCT 1878  
1621 TTCCGCGATGCAACCAAGCTTTGCTTGGACCCGAGTCCGCTCATTTCCGACTGAAGCT 1680  
1879 GCGGACAGGCGCTCGCAAAAGCTGACTCTCTGAGCTTTTCTCCGTGAGAGCTTGGGCGGC 1938  
1681 GCGGACAGGCGCTCGCAAAAGCTGACTCTCTGAGCTTTTCTCCGTGAGAGCTTGGGCGGC 1740  
1939 GCGACCTAGATGATGCGGATGCTTCTCTTGGAGTCCGCTGGGACAGGCTCGACAGAG 1998  
1741 GCGACCTAGATGCGGATGCTTCTCTTGGAGTCCGCTGGGACAGGCTCGACAGAG 1800  
1999 CTGCGGAGGCGATGCGGAATGTAAACATTCAGATGCTCTCTTCCGCGGCGCAACACCGTG 2058  
1801 CTGCGGAGGCGATGCGGAATGTAAACATTCAGATGCTCTCTTCCGCGGCGCAACACCGTG 1860  
2059 GGTATACACCCGCTACCGAGACTCCGCTCTGCGGCGGTTTGTAAAGAGCTGCCAGCTCC 2118  
1861 GGTATACACCCGCTACCGAGACTCCGCTCTGCGGCGGTTTGTAAAGAGCTGCCAGCTCC 1920  
2119 GCGGTGAGCATCTTCCGCTCTTCCGAGCGGCTTAAACAGCTCTCCAGATGCGTCCAGCA 2178  
1921 GCGGTGAGCATCTTCCGCTCTTCCGAGCGGCTTAAACAGCTCTCCAGATGCGTCCAGCA 1980  
2179 ATGAGCGAGTCTGAGAGACCAACCGGCTAGCCGAGTGGCTATGCTTATTTGCT 2238  
1981 ATGAGCGAGTCTGAGAGACCAACCGGCTAGCCGAGTGGCTATGCTTATTTGCT 2040  
2239 GATCTCTCTGATCAAAATGAAAGCTCTACACCTCGATTTACTACCTAAAGATGGCAGAG 2298  
2041 GATCTCTCTGATCAAAATGAAAGCTCTACACCTCGATTTACTACCTAAAGATGGCAGAG 2100  
2299 GAGATCTCTGATCTGCGGCTCTCATCTTGGCGCATTAAGGATATGCTGCTGCTTCCG 2358  
2101 GAGATCTCTGATCTGCGGCTCTCATCTTGGCGCATTAAGGATATGCTGCTGCTTCCG 2160  
2359 CCAGCTGCGGTAAACAGCTGGTCAACCGCTGCGCGCTGAATTCGATCTGCGAGTGCAC 2418  
2161 CCAGCTGCGGTAAACAGCTGGTCAACCGCTGCGCGCTGAATTCGATCTGCGAGTGCAC 2220  
2419 GTGCACACCCAGCACTGCGGCTGCGAGTGGCAACCTTACTTTGCTGCGAGTCAAGCT 2478  
2221 GTGCACACCCAGCACTGCGGCTGCGAGTGGCAACCTTACTTTGCTGCGAGTCAAGCT 2280  
2479 GGTGCAGATGCTGCTGAGCGTCTCCGACACCTGCTGCGCACCACTCCCGAGCCATCC 2538  
2281 GGTGCAGATGCTGCTGAGCGTCTCCGACACCTGCTGCGCACCACTCCCGAGCCATCC 2340  
2539 CTGCTGCCATTTGCTGCTGATTCGCGCACACCGCTGCGGATACCGGTTGAGCGCTCGAG 2598  
2341 CTGCTGCCATTTGCTGCTGATTCGCGCACACCGCTGCGGATACCGGTTGAGCGCTCGAG 2400  
2599 GCTGTTTCTGACCTCGAGCGTACTGGAGAGTGGCGGACTGTGCTGCTGCTGCTGCTGCTG 2658  
2401 GCTGTTTCTGACCTCGAGCGTACTGGAGAGTGGCGGACTGTGCTGCTGCTGCTGCTGCTG 2460  
2659 TCTGGAACCCAGCGCCAAACCGCTGCGCTTACCGCACGAAATCCAGCGGACAGTTG 2718  
2461 TCTGGAACCCAGCGCCAAACCGCTGCGCTTACCGCACGAAATCCAGCGGACAGTTG 2520  
2719 TCCAACTGCTGCTGAGAGCGCAACCGCTGCGCTTACCGCACGAAATCCAGCGGACAGTTG 2778  
2521 TCCAACTGCTGCTGAGAGCGCAACCGCTGCGCTTACCGCACGAAATCCAGCGGACAGTTG 2580  
2779 GACAACTACGCGCGCTTAATGAGATGCTGGGAGCGCCCAACCAAGGTCACCCCATCTCC 2838

Db 2581 GACAACTACGACGCGTATATGATGCTGGAGCGCCCAACCAAGGTCAACCCATCCTCC 2640  
Qy AAGGTTCTGTGGCGACCTCGCACTCCACTCGTTGGTGGCGGTGTGGATCCAGCAGACTTT 2898  
Db AAGGTTCTGTGGCGACCTCGCACTCCACTCGTTGGTGGCGGTGTGGATCCAGCAGACTTT 2700  
Qy GCTGCCGATCCAAAGATGACATCCAGACTCTGTCTATCCGTTCTCTGCGCGGAG 2958  
Db GCTGCCGATCCAAAGATGACATCCAGACTCTGTCTATCCGTTCTCTGCGCGGAG 2760  
Qy CTTGGTAACTCCAGCTGGCTGGCCAGAGCACTCGCGCAACCGCGCACTGGAGGCCCGC 3018  
Db CTTGGTAACTCCAGCTGGCTGGCCAGAGCACTCGCGCAACCGCGCACTGGAGGCCCGC 2820  
Qy TCCGAAGGCAAGGCACTCTGACCGAAGTTCTTGAAGAGAGAGAGCGGCACTCGACGCT 3078  
Db TCCGAAGGCAAGGCACTCTGACCGAAGTTCTTGAAGAGAGAGAGCGGCACTCGACGCT 2880  
Qy GATGATTCGAAGGCACTGCGCAATAGCTCAACCGCTGCTTCCCGAAGCCAAACCGAA 3138  
Db GATGATTCGAAGGCACTGCGCAATAGCTCAACCGCTGCTTCCCGAAGCCAAACCGAA 2940  
Qy GAGTTCTCGAGCAACCGTCCGCGCTTCGCGCAACCACTCTCGCTGGATGATCGTGAATC 3198  
Db GAGTTCTCGAGCAACCGTCCGCGCTTCGCGCAACCACTCTCGCTGGATGATCGTGAATC 3000  
Qy TTCTACGCTGCTGCGAGCGCGGAGACTTTGATCCGCTGCGCAGATGCGCACCCCA 3258  
Db TTCTACGCTGCTGCGAGCGCGGAGACTTTGATCCGCTGCGCAGATGCGCACCCCA 3060  
Qy TTTCTACGCTGCTGCGAGCGCGGAGACTTTGATCCGCTGCGCAGATGCGCACCCCA 3318  
Db TTTCTACGCTGCTGCGAGCGCGGAGACTTTGATCCGCTGCGCAGATGCGCACCCCA 3120  
Qy CTGCTTGTTCGCTGATGCGATCTCTGAGCGAGCAAGGATGCGCAATGTTGTG 3378  
Db CTGCTTGTTCGCTGATGCGATCTCTGAGCGAGCAAGGATGCGCAATGTTGTG 3378  
Qy GCGAAGTCAAGGCACTGCGCAATAGCTCAACCGCTGCTTCCCGAAGCCAAACCGAA 3438  
Db GCGAAGTCAAGGCACTGCGCAATAGCTCAACCGCTGCTTCCCGAAGCCAAACCGAA 3240  
Qy GGTGTTGTCAGGCTGCTGAGGCTGATGAGTCAAGGCTGAGGCTGAGGCTGAGTCCGA 3498  
Db GGTGTTGTCAGGCTGCTGAGGCTGATGAGTCAAGGCTGAGGCTGAGGCTGAGTCCGA 3300  
Qy GGTGTTGTCAGGCTGCTGAGGCTGATGAGTCAAGGCTGAGGCTGAGGCTGAGTCCGA 3558  
Db GGTGTTGTCAGGCTGCTGAGGCTGATGAGTCAAGGCTGAGGCTGAGGCTGAGTCCGA 3360  
Qy GCGTTGTTGTTCTGCTGCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGTCCGA 3618  
Db GCGTTGTTGTTCTGCTGCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGTCCGA 3420

RESULT 4  
US-09-974-973-1  
; Sequence 1, Application US/09974973  
; Patent No. US20020177202A1  
; GENERAL INFORMATION:  
; APPLICANT: Hanke, Paul D.  
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacterium  
; FILE REFERENCE: 1533, 1230001/MAC/RGM  
; CURRENT APPLICATION NUMBER: US/09/974, 973  
; CURRENT FILING DATE: 2001-10-21  
; PRIOR APPLICATION NUMBER: US 60/239, 913  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 3474  
; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(3474)  
US-09-974-973-1  
Query Match 93.9%; Score 3398.8; DB 9; Length 3474;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
Qy 148 GTGACTGTATCAACCCCTTGGCGGCTCTCTTGTGAAAGGATTAATCTCTAGTCTGCACT 207  
Db 1 GTGACTGTATCAACCCCTTGGCGGCTCTCTTGTGAAAGGATTAATCTCTAGTCTGCACT 60  
Qy 208 CACACATCTTCAACGCTTCCAGCACTTCAAAAGATCTTGGTAGCAAAACCGCGCGGAATC 267  
Db 61 CACACATCTTCAACGCTTCCAGCACTTCAAAAGATCTTGGTAGCAAAACCGCGCGGAATC 120  
Qy 268 GCGGTCCTGCTTCCGTCGAGCACTCGAAACCGGTGAGCCAGCTGATTTACCCC 327  
Db 121 GCGGTCCTGCTTCCGTCGAGCACTCGAAACCGGTGAGCCAGCTGATTTACCCC 180  
Qy 328 CGTGAAGATCGGGATCAATCCACCGCTCTTTTCTTCTGAAGCTGTCGCAATGGTACC 387  
Db 181 CGTGAAGATCGGGATCAATCCACCGCTCTTTTCTTCTGAAGCTGTCGCAATGGTACC 240  
Qy 388 GAAGGCTCACAAGTCAAGCGGTACCTGACATCGATGAAATATCGGTGCAAGCTTAAAAA 447  
Db 241 GAAGGCTCACAAGTCAAGCGGTACCTGACATCGATGAAATATCGGTGCAAGCTTAAAAA 300  
Qy 448 GTTAAAGCAGATGCCATTTACCGGGATACCGCTTCTCTGAAATGCGCAAGCTTGGC 507  
Db 301 GTTAAAGCAGATGCCATTTACCGGGATATGGCTTCTCTGAAATGCGCAAGCTTGGC 360  
Qy 508 CGCGAGTGTGCGGAAAAAGGCAATTAATTTATGCGCAACCCAGAGGTTCTTGAATC 567  
Db 361 CGCGAGTGTGCGGAAAAAGGCAATTAATTTATGCGCAACCCAGAGGTTCTTGAATC 420  
Qy 568 ACCGTTGATAGTCTCGCGGTACCGCGGTACCGCTTCTCTGAAATGCGCAAGCTTGGC 627  
Db 421 ACCGTTGATAGTCTCGCGGTACCGCGGTACCGCTTCTCTGAAATGCGCAAGCTTGGC 480  
Qy 628 GAATCCACCCCGAGCAAAACATCGATGAGATCGTTTAAAGCGCTGAAAGCGCACTTAC 687  
Db 481 GAATCCACCCCGAGCAAAACATCGATGAGATCGTTTAAAGCGCTGAAAGCGCACTTAC 540  
Qy 688 CCAATCTTTGTGAGGCACTGCGCGGTGCTGCGGAGCGGCTGATGCTTCTGCTTCA 747  
Db 541 CCAATCTTTGTGAGGCACTGCGCGGTGCTGCGGAGCGGCTGATGCTTCTGCTTCA 600  
Qy 748 CCGTATGAGCTTCGCAAAATAGCAAGAGCATCTCGTGAAGCTGAAAGCGCTTTCGGC 807  
Db 601 CCGTATGAGCTTCGCAAAATAGCAAGAGCATCTCGTGAAGCTGAAAGCGCTTTCGGC 660  
Qy 808 GATGGCGGCTATATGTCGAAAGCTGTCGAGTAAACCTCAGCATTTGAAGTGCAGATC 867  
Db 661 GATGGCGGCTATATGTCGAAAGCTGTCGAGTAAACCTCAGCATTTGAAGTGCAGATC 720  
Qy 868 CTTGGCGATCACACTCGAGAGTTGTACACCTTTATGAAGCTGACTGCTCACTGCAAGCT 927  
Db 721 CTTGGCGATCACACTCGAGAGTTGTACACCTTTATGAAGCTGACTGCTCACTGCAAGCT 780  
Qy 928 CGTCACCAAAAGTTCTCGAAATTCGCCAGCAGCATTTGGATCCAGAACTGCTGTAT 987  
Db 781 CGTCACCAAAAGTTCTCGAAATTCGCCAGCAGCATTTGGATCCAGAACTGCTGTAT 840  
Qy 988 CGCATTTGTGGGATGAGTAAAGTTCTCGCGCTCCATTTGTTTACCGAGCGCGGGAAC 1047  
Db 841 CGCATTTGTGGGATGAGTAAAGTTCTCGCGCTCCATTTGTTTACCGAGCGCGGGAAC 900  
Qy 1048 GTGGAATTTCTGCTCGATGAAAGGCAACCACTCTTTCATCGAAATGAACCACTATC 1107  
Db 901 GTGGAATTTCTGCTCGATGAAAGGCAACCACTCTTTCATCGAAATGAACCACTATC 960

1108 CAGGTTGAGCA CACCGTGA CTGAAGA GTCAACCGAGGTGGACCTGGTGAAGGCGCAGATG 1167  
Db CAGGTTGAGCA CACCGTGA CTGAAGA GTCAACCGAGGTGGACCTGGTGAAGGCGCAGATG 1020  
1168 CGCTTGGCTGCTGGTGCACCTTTGAAGAAATTGGGTCTTGACCCAGATAAGATCAAGACC 1227  
Db CGCTTGGCTGCTGGTGCACCTTTGAAGAAATTGGGTCTTGACCCAGATAAGATCAAGACC 1080  
1228 CACGCTGACGCA CTGCACTGCGCATCA CACGGAAGATCCAAACAACCGCTTCGGCCCA 1287  
Db CACGCTGACGCA CTGCACTGCGCATCA CACGGAAGATCCAAACAACCGCTTCGGCCCA 1140  
1288 GATACCGGAATCATCA CCGGTACCGCTCACGACGCGAGCTGGCGTTCTGCTTGAAGGT 1347  
Db GATACCGGAATCATCA CCGGTACCGCTCACGACGCGAGCTGGCGTTCTGCTTGAAGGT 1200  
1348 GCAGCTCAGCTCGGTGGCGAAATCA CCGCACTTTGATCTCATCTGGTGAAGATGACC 1407  
Db GCAGCTCAGCTCGGTGGCGAAATCA CCGCACTTTGATCTCATCTGGTGAAGATGACC 1260  
1408 TGCGGTGTTCCGACTTCGAACTGCTGTTGCTCGTGACACGCGCGTGGCTGCTGAGTTC 1467  
Db TGCGGTGTTCCGACTTCGAACTGCTGTTGCTCGTGACACGCGCGTGGCTGCTGAGTTC 1320  
1468 ACCGTGCTGCTGTTGCAACCAATTTGTTTCTTGGTGGCTTCTGCGTGGCTTCTGCGGGAAGAAC 1527  
Db ACCGTGCTGCTGTTGCAACCAATTTGTTTCTTGGTGGCTTCTGCGTGGCTTCTGCGGGAAGAAC 1380  
1528 TTCACTTCCAGCGCATCGCCACCGATTCANTGCCGATCA CCGCACCCTCCCTCAGGCT 1587  
Db TTCACTTCCAGCGCATCGCCACCGATTCANTGCCGATCA CCGCACCCTCCCTCAGGCT 1440  
1588 CCACTGCTGATGATGAGCAGGACGCACTCTGATTTACTTGGCAGATGTCA CCGTGAAC 1647  
Db CCACTGCTGATGATGAGCAGGACGCACTCTGATTTACTTGGCAGATGTCA CCGTGAAC 1500  
1648 AAGCTCATGCTGGTCCAGAGATGTGAGCTCTTATCGATAAGCTGCTTAACATC 1707  
Db AAGCTCATGCTGGTCCAGAGATGTGAGCTCTTATCGATAAGCTGCTTAACATC 1560  
1708 AAGGATCTGCACTGCGACGCGTTCCTGCTGACCGCTGAAAGCAGCTTGGCCACGCGG 1767  
Db AAGGATCTGCACTGCGACGCGTTCCTGCTGACCGCTGAAAGCAGCTTGGCCACGCGG 1620  
1768 TTTGCTCGTGA TCTCGTGAGCAGGACGCACTGGCAGTTACTGATACCACTTCCCGAT 1827  
Db TTTGCTCGTGA TCTCGTGAGCAGGACGCACTGGCAGTTACTGATACCACTTCCCGAT 1680  
1828 GCACACAGTCTTTGCTTGCGACCCGAGTCCGCTCATTTGGCACTGAAGCTTCGGGACAG 1887  
Db GCACACAGTCTTTGCTTGCGACCCGAGTCCGCTCATTTGGCACTGAAGCTTCGGGACAG 1740  
1888 GCGGTGCAAGCTGACTCTGAGTTTGTCCGTGAGAGGCTGGGCGGCGGACCTTAC 1947  
Db GCGGTGCAAGCTGACTCTGAGTTTGTCCGTGAGAGGCTGGGCGGCGGACCTTAC 1800  
1948 GATGTGCGGATCGTTTCTCTTTGAGATCCGTGGGACAGGCTCGACGAGCTTGGCGGAG 2007  
Db GATGTGCGGATCGTTTCTCTTTGAGATCCGTGGGACAGGCTCGACGAGCTTGGCGGAG 1860  
2008 GCGATGCCGAATGTAACATTCAGTCTGCTGCGGCGGCGAACA CCGTGGGATACACC 2067  
Db GCGATGCCGAATGTAACATTCAGTCTGCTGCGGCGGCGAACA CCGTGGGATACACC 1920  
2068 CCGTACCCAGACTCCGCTGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGCGGTGGAC 2127  
Db CCGTACCCAGACTCCGCTGCGCGCTTTGTTAAGGAAGCTGCCAGCTCCGCGGTGGAC 1980  
2128 ATCTTCGCGATCTTCGACGCGCTTAA CAGCGTCTCCAGATGGCTCAGCAATTCAGCGCA 2187  
Db ATCTTCGCGATCTTCGACGCGCTTAA CAGCGTCTCCAGATGGCTCAGCAATTCAGCGCA 2040

2188 GTCTTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2247  
Db GTCTTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2100  
2248 GATCCAAATGAAAAGCTCTACACCTTGGAATTA CTAAAGATGGCAGAGAGATCGTC 2307  
Db GATCCAAATGAAAAGCTCTACACCTTGGAATTA CTAAAGATGGCAGAGAGATCGTC 2160  
2308 AGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGTGTGCTCTTCGCCAGCTGG 2367  
Db AGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGTGTGCTCTTCGCCAGCTGG 2220  
2368 GTAAACCAAGCTGCTCA CCGCACTGCGCGCTGAATTCGATCTGCCAGTGCAGATC 2427  
Db GTAAACCAAGCTGCTCA CCGCACTGCGCGCTGAATTCGATCTGCCAGTGCAGATC 2280  
2428 CAGCACTGCGGCTGGCGCAGCTGGCAACTTCTTGTGTCAGCTCAAGCTGGTGAGAT 2487  
Db CAGCACTGCGGCTGGCGCAGCTGGCAACTTCTTGTGTCAGCTCAAGCTGGTGAGAT 2340  
2488 GCTGTTGACGCTGCTTCCGACCACTGTCTGGCA CCACTCCGAGCCATCCCTGTCTGCC 2547  
Db GCTGTTGACGCTGCTTCCGACCACTGTCTGGCA CCACTCCGAGCCATCCCTGTCTGCC 2400  
2548 ATTGTTGCTGATTCGGGCA CCACTCCGCTCGGATACCGGTTTGAGCTCGAGCTGTTCT 2607  
Db ATTGTTGCTGATTCGGGCA CCACTCCGCTCGGATACCGGTTTGAGCTCGAGCTGTTCT 2460  
2608 GACCTCGAGCGTACTTGGGAAGCAGTGC CGGACTGTACTCTGCACTTTGAGTCTGGAAC 2667  
Db GACCTCGAGCGTACTTGGGAAGCAGTGC CGGACTGTACTCTGCACTTTGAGTCTGGAAC 2520  
2668 CAGGCCCCA CCGGTGCTGTACCGCCAGCAATCC CAGGCGGACAGTTGTCACCTG 2727  
Db CAGGCCCCA CCGGTGCTGTACCGCCAGCAATCC CAGGCGGACAGTTGTCACCTG 2580  
2728 CGTSCACAGGCGCA CCGCACTGGGCTTGGGATCG GTTTCGAACTCATCGAAGACAATAC 2787  
Db CGTSCACAGGCGCA CCGCACTGGGCTTGGGATCG GTTTCGAACTCATCGAAGACAATAC 2640  
2788 GCAGCCCTTAATGAGATGTGGGACCGCCNA CCAAGTCA CCCCCTCCAGAGTTGTT 2847  
Db GCAGCCCTTAATGAGATGTGGGACCGCCNA CCAAGTCA CCCCCTCCAGAGTTGTT 2700  
2848 GGCACCTCGCACTCCACCTCGTTGGTGGGCTGTGGATCCAGCAGACTTTGTGCGGAT 2907  
Db GGCACCTCGCACTCCACCTCGTTGGTGGGCTGTGGATCCAGCAGACTTTGTGCGGAT 2760  
2908 CCAAAAAGTACGACATCC CAGACTCTGTCA TCGGTTCTGCGGCGGAGCTTGTAAAC 2967  
Db CCAAAAAGTACGACATCC CAGACTCTGTCA TCGGTTCTGCGGCGGAGCTTGTAAAC 2820  
2968 CCTCCAGTGGCTGGCCAGAGCCACTCGC GACCGCGCACTGGAAGCCGCTCCGAAGGC 3027  
Db CCTCCAGTGGCTGGCCAGAGCCACTCGC GACCGCGCACTGGAAGCCGCTCCGAAGGC 2880  
3028 AAGGCACTCTGACGGAAGTTCTGAGAAAGAGCAGCGGCA CCGTGCAGCGTGTATTC 3087  
Db AAGGCACTCTGACGGAAGTTCTGAGAAAGAGCAGCGGCA CCGTGCAGCGTGTATTC 2940  
3088 AAGGAAGTCCGAATAGCTCAACCGCTGTGTTTCCGGAAGCCAA CCGAGAGTTCTCTC 3147  
Db AAGGAAGTCCGAATAGCTCAACCGCTGTGTTTCCGGAAGCCAA CCGAGAGTTCTCTC 3000  
3148 GAGCACCTGCCCTTCGGCAACACTCTGCGTGGATGATTCGTGAATTTCTTCTAGGCG 3207  
Db GAGCACCTGCCCTTCGGCAACACTCTGCGTGGATGATTCGTGAATTTCTTCTAGGCG 3060  
3208 CTGTCGAGGCGCGGAGACTTTGATCCGCTGCCAGATGTGCGCACCCCACTCTTGT 3267  
Db CTGTCGAGGCGCGGAGACTTTGATCCGCTGCCAGATGTGCGCACCCCACTCTTGT 3120  
3268 CGCTGGATGCGATCTCTGAGCCAGACGATGAAGGATATGGCAATGTGTGGCCAAAGTC 3327



1441 CCACCTGCGGATGATGAGCAGGAGCGCATCTCTGGATTACTTTGGCAGATGTCACCGTGAAC 1500  
1648 AAGCCTCATGCTGCTGCTCCAAAGAGATGTTGACGATCTCTATCGATAAGCTCCCTAAACATC 1707  
1501 AAGCCTCATGCTGCTGCTCCAAAGAGATGTTGACGACCAACATCGATAAGCTGCCCAATC 1560  
1708 AAGGATCTGCCACTGCGCAGCGCGTTCCCGTGACCGCCTGAAGCAGCTTGGGCCAGCGCG 1767  
1561 AAGGATCTGCCACTGCGCAGCGCGTTCCCGTGACCGCCTGAAGCAGCTTGGGCCAGCGCG 1620  
1768 TTTGCTCGTGATCTCCGTTGAGCAGACGACCTGGCAGTACTGTATACCACTTCCGCGAT 1827  
1621 TTTGCTCGTGATCTCCGTTGAGCAGACGACCTGGCAGTACTGTATACCACTTCCGCGAT 1680  
1828 GCACACCACTTTGCTTGGCAGACCGAGTCCGCTCATTTGCACTGAAGCCTTGGCGCAGAG 1887  
1681 GCACACCACTTTGCTTGGCAGACCGAGTCCGCTCATTTGCACTGAAGCCTTGGCGCAGAG 1740  
1888 GCGCTGCGAAAGCTGACTCTCAGCTTTTGTCCGTTGAGGAGCTGGGGCGGCGCACCTTAC 1947  
1741 GCGCTGCGAAAGCTGACTCTCAGCTTTTGTCCGTTGAGGAGCTGGGGCGGCGCACCTTAC 1800  
1948 GATGTGCGGATCGTTTCTCTTTGAGATCCGTTGGGACAGGCTGACGAGCTGCGGAG 2007  
1801 GATGTGCGGATCGTTTCTCTTTGAGATCCGTTGGGACAGGCTGACGAGCTGCGGAG 1860  
2008 GCGATGCGGAATGTAACATTCAGATGCTGCTTCCGCGCGCAACACCGTGGGATACACC 2067  
1861 GCGATGCGGAATGTAACATTCAGATGCTGCTTCCGCGCGCAACACCGTGGGATACACC 1920  
2068 GCGTACCCAGACTCGTCTGCGCGCTTTGTTAAGGAGCTGCGCAGCTCCGCGGTGGAC 2127  
1921 GCGTACCCAGACTCGTCTGCGCGCTTTGTTAAGGAGCTGCGCAGCTCCGCGGTGGAC 1980  
2128 ATCTTCGCGATCTTGCAGCGCTTAAACAGCTCTCCGAGTCCGTCAGCAATCGACGCA 2187  
1981 ATCTTCGCGATCTTGCAGCGCTTAAACAGCTCTCCGAGTCCGTCAGCAATCGACGCA 2040  
2188 GTCTGAGACCAACACCGCGGTAGCCAGTGGCTATGGCTTAATCTGGTGATCTCTCT 2247  
2041 GTCTGAGACCAACACCGCGGTAGCCAGTGGCTATGGCTTAATCTGGTGATCTCTCT 2100  
2248 GATCCAAATGAAGCTCTACACCTGGATTAACCTTAAGATGGCAGAGGATCGTC 2307  
2101 GATCCAAATGAAGCTCTACACCTGGATTAACCTTAAGATGGCAGAGGATCGTC 2160  
2308 AAGTCTGGCGCTCACAATCTGGCCATTAAGGATATGGCTGCTCTGCTTCCGCCAGCTGCG 2367  
2161 AAGTCTGGCGCTCACAATCTGGCCATTAAGGATATGGCTGCTCTGCTTCCGCCAGCTGCG 2220  
2368 GTAAACCAAGCTGGTACCCGCACTGCGCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2427  
2221 GTAAACCAAGCTGGTACCCGCACTGCGCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2280  
2428 CACGACACTCGGGTGGCCAGCTGCTTGTGCTGAGCTCAAGCTGGTGCGAGAT 2487  
2281 CACGACACTCGGGTGGCCAGCTGCTTGTGCTGAGCTCAAGCTGGTGCGAGAT 2340  
2488 GCTGTTGACGGTCTTCCGCAACAATGCTGCGCAACAATCTCCAGCCATCCCTGTCTGCC 2547  
2341 GCTGTTGACGGTCTTCCGCAACAATGCTGCGCAACAATCTCCAGCCATCCCTGTCTGCC 2400  
2548 ATTGTTGCTGCAATTCGCGCACACCGCTGCGGATACCGGTTTTCGAGCTCGAGCTGTTTCT 2607  
2401 ATTGTTGCTGCAATTCGCGCACACCGCTGCGGATACCGGTTTTCGAGCTCGAGCTGTTTCT 2460  
2608 GACCTCGACCGTACTGCGGAAGAGTGGCGGACTGATCTGCTGCAATTTGAGTCTGGAACC 2667  
2461 GACCTCGACCGTACTGCGGAAGAGTGGCGGACTGATCTGCTGCAATTTGAGTCTGGAACC 2520  
2668 CCAGGCCCCAACCGCTCGCTCTACCGCCACGAAATCCAGGCGGAGAGTGTGCCAACCTG 2727  
2521 CCAGGCCCCAACCGCTCGCTCTACCGCCACGAAATCCAGGCGGAGAGTGTGTCCAACCTG 2580

2728 CGTGCACAGGCGCACCGCACTGGGCGCTTGGGATCGTTTGGAACTCATCAAGACAACTAC 2787  
2581 CGTGCACAGGCGCACCGCACTGGGCGCTTGGGATCGTTTGGAACTCATCAAGACAACTAC 2640  
2788 GCAGCGGTTAATGAGATGCTGGAGCGCCCAACCAAGGTCAACCCATCTCCAAAGTGTGT 2847  
2641 GCAGCGGTTAATGAGATGCTGGAGCGCCCAACCAAGGTCAACCCATCTCCAAAGTGTGT 2700  
2848 GCGGACCTCGCACTCCACCTCGTTGGTGGGCTGTGGATCCAGCAGACTTTGCTCCGAT 2907  
2701 GCGGACCTCGCACTCCACCTCGTTGGTGGGCTGTGGATCCAGCAGACTTTGCTCCGAT 2760  
2908 CCACAAAGTACGACATCCAGACTCTGTATCGCGTTCCTTGGCGGAGGCTTGGTAAC 2967  
2761 CCACAAAGTACGACATCCAGACTCTGTATCGCGTTCCTTGGCGGAGGCTTGGTAAC 2820  
2968 CTTCCAGGTGGCTGGCGAGGCACTTGGCGCAACCCGCGCACTGGAAAGGCGCTTCCGAAAGC 3027  
2821 CTTCCAGGTGGCTGGCGAGGCACTTGGCGCAACCCGCGCACTGGAAAGGCGCTTCCGAAAGC 2880  
3028 AAGGACCTCTGACGGAAGTTCCTGAGGAGAGCAGGCGCACTCGAGCTGATGATTC 3087  
2881 AAGGACCTCTGACGGAAGTTCCTGAGGAGAGCAGGCGCACTCGAGCTGATGATTC 2940  
3088 AAGGAACTGCGCAATAGCTCAACCGCTTGTTCCTGAAAGCCAAACCGAAGTTCCTC 3147  
2941 AAGGAACTGCGCAATAGCTCAACCGCTTGTTCCTGAAAGCCAAACCGAAGTTCCTC 3000  
3148 GAGCACCTGCGCGCTTGGGCAACACTCTGCGCTGGATGATCGTGAATTCCTTACGCG 3207  
3001 GAGCACCTGCGCGCTTGGGCAACACTCTGCGCTGGATGATCGTGAATTCCTTACGCG 3060  
3208 CTGGTCGAAAGGCGCGGAGACTTTGATCCGCTTGCAGATGTGCGCACCCCACTGTTGT 3267  
3061 CTGGTCGAAAGGCGCGGAGACTTTGATCCGCTTGCAGATGTGCGCACCCCACTGTTGT 3120  
3268 CGCTCGATGCGATCTCTGAGCAGACGATAGGATGCGCAATGTTGTCGCGCAACGTC 3327  
3121 CGCTCGATGCGATCTCTGAGCAGACGATAGGATGCGCAATGTTGTCGCGCAACGTC 3180  
3328 AAGCGCCAGATCCGCCAATGCTGCGTGGTCCGCTCGTTCGTTGAGTCTGTACCCGCAAC 3387  
3181 AAGCGCCAGATCCGCCAATGCTGCGTGGTCCGCTCGTTCGTTGAGTCTGTACCCGCAAC 3240  
3388 GCAGAAAGGCGAGATTCCTCCAAAGGCGCATGTTGTCGACCAATTCGCTGGTGTGTC 3447  
3241 GCAGAAAGGCGAGATTCCTCCAAAGGCGCATGTTGTCGACCAATTCGCTGGTGTGTC 3300  
3448 ACCGTGACTGTTGCTGAAGTGTGAGTCAAGGCTGGAGATGAGTGCATCATCTCGAG 3507  
3301 ACTGTGACTGTTGCTGAAGTGTGAGTCAAGGCTGGAGATGAGTGCATCATCTCGAG 3360  
3508 GCTATGAAGATGGAACAACTCATCTGCTTCTGTTGACGCGCAAAATCGATCGGTTGTG 3567  
3361 GCTATGAAGATGGAACAACTCATCTGCTTCTGTTGACGCGCAAGATTGAACGCGTTGTG 3420  
3568 GTTCTCTGCTGACGAGAGTGGAGTGGGAGCTTGTGATCGTCTGTTTCTCTAA 3621  
3421 GTTCTCTGCTGACGAGAGTGGAGTGGGAGCTTGTGATCGTCTGTTTCTCTAA 3474

## RESULT 6

US-10-282-122A-17701  
; Sequence 17701, Application US/10282122A  
; Publication No. US20040029129A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Olsen, Kari  
; APPLICANT: Zyskind, Judith



APPLICANT:	Wall, Daniel	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	TITLE OF INVENTION:	Identification of Essential Genes in Microorganisms
APPLICANT:	Trawick, John	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	FILE REFERENCE:	ELIUPA 034A
APPLICANT:	Carr, Grant	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	CURRENT APPLICATION NUMBER:	US/10/282,122A
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	CURRENT FILING DATE:	2003-02-20
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR APPLICATION NUMBER:	60/191,078
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR FILING DATE:	2000-03-21
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR APPLICATION NUMBER:	60/206,848
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR FILING DATE:	2000-05-23
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR APPLICATION NUMBER:	60/207,727
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR FILING DATE:	2000-05-26
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR APPLICATION NUMBER:	60/230,335
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR FILING DATE:	2000-09-06
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR APPLICATION NUMBER:	60/230,347
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR FILING DATE:	2000-09-09
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR APPLICATION NUMBER:	60/242,578
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR FILING DATE:	2000-10-23
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR APPLICATION NUMBER:	60/253,625
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR FILING DATE:	2000-11-27
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR APPLICATION NUMBER:	60/257,931
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR FILING DATE:	2000-12-22
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR APPLICATION NUMBER:	60/257,636
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR FILING DATE:	2001-02-09
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR APPLICATION NUMBER:	60/269,308
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	PRIOR FILING DATE:	2001-02-16
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	Remaining prior application data removed - See File Wrapper or PALM.	
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	NUMBER OF SEQ ID NOS:	78614
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	SOFTWARE:	Patentin version 3.1
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	SEQ ID NO	17701
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	LENGTH:	3423
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	TYPE:	DNA
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	ORGANISM:	Corynebacterium diptheriae
APPLICANT:	Yamamoto, Robert	APPLICANT:	Yamamoto, Robert	APPLICANT:	Xu, H.	US-10-282-122A-17701	
Query Match	53.2%	Score	1927.4	DB 12	Length	3423	
Best Local Similarity	73.4%	Pred. No. 0					
Matches 2494	Conservative	0	Mismatches	856	Indels	9	Gaps 2
Qy	213	ATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAAACCGCGCGGAGAAATCGCGT	272				
Db	27	AACCTCTACGGTAAATCCATTTGTCAAAGATCTCTCGTTGCTAACCGTGGCAGATCGCGT	86				
Qy	273	CGGTGCTTTCCGTGCAGCAGCTCGAAAACGGGTGACGACCGTAGCTATTTTACCCCGCTGA	332				
Db	87	CGGAGCTTTTCGTGCAGCCTTTGAAAACGGGTGCTGCCACGGTAGCTCTTACCCGACGA	146				
Qy	333	AGATCGGGGATCATATCCACCGCTCTTTTGTCTTGAAGCTGTCCGATTTGGTACCGAAGG	392				
Db	147	AGACCGGTAAATTCATTCACCGCATCGTTGGCTCTTGAAGCTGTCTCATTTGGTGAGGCGG	206				
Qy	393	CTCACAGTCAAGCGGTACCTTGGACATCGATGAATATTCGGTGCAGCTAAAAAGTTAA	452				
Db	207	ATCTGCGGTCAAAGCGTACCTCGATTCGACGAATCATCGTCTGCTAAACAAACGGG	266				
Qy	453	AGCAGATGCCATTTACCGGGATACGGCTTCTCTGTGAATAATGCCACGCTTGGCCGGA	512				
Db	267	TGCCGACGCAATTTACCCAGCTACGGCTTCCCTTCCGAAAACGCCAGCTTGTCTCGTA	326				
Qy	513	GTGTGCGGAAAACGGCATTAATCTTTATTTGGCCCAACCCAGAGGTTCTTGATCTCACCGG	572				
Db	327	GTGCGCTGAAAACGGCATAACCTTTATCGGTCCACCTCCATCAGTGTGGAGCTTACCGG	386				
Qy	573	TGATAAGTCTCGCGGTAAACCGCGGAGAGAGCTGGTCTGCCAGTTTGGCGGATC	632				
Db	387	TGATAAAGCAGCTGCTGTGTTACCGCAGCACGCGAGGAGGTTTGGCCACGCTGATCGAGC	446				
Qy	633	CACCCCGACGAAAACATTCGATGAGATCGTTTAAAAAGCGCTCAAGGCCAGATTTACCCCAT	692				



1773 TCGTGATCTCCGTGAGCAGCAGCTCGCAGTTACTGATACCACTTTCGCGATGACCA 1832  
1833 CGAGTCTTTGCTGCGACCGAGTCCGCTCAATCGCACTGAAGCTCGCGGAGAGGCGGT 1892  
1644 TCAGTCGCTTCTCGCAACCCGCGTGCCTCAACACACGCTGATCGACGAGCAGCGACGT 1703  
1893 CGCAAGCTGACTCTGAGCTTTTCTCGTGGAGGCTCGCGGCGCGGAGCTTACGATGT 1952  
1704 CGCAAGCTCACCCAGAGCTGCTCTCTGTAGAGCATGGGCTGCTACTATGACGT 1763  
1953 GCGGATGCTTCTCTTTTGAAGATTCCTGTGGACAGGCTCGACGAGTTCGCGAGGCGAT 2012  
1764 TCGTATGCGCTTCTTACATGAGACCCCTCGGAGCGTCTCGACCACTTCGCTGAGGCGAT 1823  
2013 GCGGATGTAACATTTAGATCTCTTTCGCGGCGCAACCGTGGGATACACCCGTA 2072  
1824 GCGCAAGCTCAATTTAGATGCTCTGTGGAGGCGCAACAGTTCGCTACACCGTA 1883  
2073 CCGAGACTCCGCTCTCGCGCGGCTTTTGAAGAGCTGCCAGCTCCGCGGTGGACATCTT 2132  
1884 CCGAGATTTCTGATGCAAGGCTTTTGTGCAAGAGCCGACGCTCTGCTGTGTGATATCTT 1943  
2133 CCGCATCTTCGAGCGCTTAACGAGCTCTCCAGATGCTCCAGCAATCGACGAGTCTCT 2192  
1944 CCGCATCTTCGAGCGCTTTAAGAGCTCTCCAGATGCTCCAGCAATCGAGCATGAGTCTCT 2003  
2193 GGAGACCAACACCGCGGTAGCGAGTGGCTATGCTTTATTTGCTGATCTCTGATCC 2252  
2004 GGAACCAACACGATTTGCGAAGTGGCAATGGCTACTCCGCTGATCTGACCAACC 2063  
2253 AATGGAAGCTCTACCGCTGATTAACCTAAGATGCGAGGAGATGCTCAAGTC 2312  
2064 AAGTGAAGAGCTCTACACTGATTAACCTTAAGCTTTCGCGAGGAATTCGTTAAGTC 2123  
2313 TGGCGCTCACATCTTGGCCATTAAGGATATGGCTGCTTCTGCTCCCGCAGCTCGGTAAC 2372  
2124 CGGTGCACATGTTCTAGCCATTAAGACATGCTGATGAAGCGCAGCGCAGCAAC 2183  
2373 CAAGCTGTCACGCACTCGCGCGTGAATTCGATCTGCGAGTGCAGTGCACCCAGCA 2432  
2184 CAAGCTGTCACGCACTACGACGCAACTTCGACCTGCGGTTCTATGTTCAACCCAGCA 2243  
2433 CACTGCGGTGGCCAGCTGGCAACTTACTTTCGCTGAGCTCAAGCTGGTGCAGATGCTGT 2492  
2244 CACTGCTGTGGCCAGCTGGCTACGTTATTTGGGCTGCAGCAGCGCTGGTGCAGATGCCGT 2303  
2493 TGACGCTGCTTCGCAACCACTGCTTGGGACCACTCCAGCCATCCCTGCTGCGCATTTGT 2552  
2304 CGATGCTGCTCCGCGCCACTGTGCGGCCACACGTCGCGAGCCATCGTTGTCTGCAATTTGT 2363  
2553 TGTGCTATTCGCGCACACCGCTCGGATACCGGTTTGAAGCTTCGAGGCTGTTTCTGACCT 2612  
2364 GGCAGATTCGCTTAACACTACCGGATACGGGTCTCTCTGAGCAGCTCGATCGAT 2423  
2613 CGAGCCGCTACTGGGAAGAGTGGCGGAGTGTACTGCTGCAATTTGATCTGGAAACCCAGG 2672  
2424 GGAACCGTACTGGGAAGAGTGGCGGAACTCTACGCACTTCGAACTCGGAACCCAGG 2483  
2673 CCCAACCGGTGCGCTTACCGCCAGCAATCCAGGCGGACAGTTGTCCAACTCGCTGTC 2732  
2484 ACCAATGCGCGCTCTACAGCAGCAATCCAGGCGGAGCTGTCCATCTCGCGCG 2543  
2733 ACAGGCCACCGCACTGGGCTTTCGGATCGTTTGAATCTATCGAAGCACTACGAGC 2792  
2544 ACAGGCCAATGCACTCGGCTTGGCAGATCGTTTCGAGCTCATCGAGGACTACTACGAGC 2603  
2793 CGTTAATGAGTCTGGGAGCGCCCAACCAAGTTCACCCCATCTCCAAAGTTGTTGGCGA 2852  
2604 CGTTAATGAGTCTGGGAGCGCCCAACCAAGTTCACCCCATCTCCAAAGTCTGTTGGCGA 2663

2853 CCTCGACTCCACCTCGTTGGTGGGCTGTGATCCAGCAGACTTTGCTGCGATCCACA 2912  
2664 TCTTGCATCTTACTAGTCGCGCAGAGATTAAACCCAGGACTTTGCGAGAGATCCCCA 2723  
2913 AAAGTACGATCCAGACTCTGTGATCGGTTCTGCGCGGAGCTTGGTAAACCTTCC 2972  
2724 AAAGTACGATATTTCTGATCTCGTCAATTCGCTTCTCTCGTGTGAGCTCGGACCCCTCC 2783  
2973 AGTGTGCTGGCCAGAGCCTACGCGCTCTGTTCGCGACCGGCTGCTGGAAGCGCTCCGAGGCAAGGC 3032  
2784 AGCGGCTGGCCAGAGAACTGCGTGTAAAGGCAATTGGCAGGACGTAAGGAATCAAGGA 2843  
3033 ACTCTGAGCGAAGTTCCTGAGAGAGAGCAGCGCACCTCGACGCTGATGATTCACAAGGA 3092  
2844 CACATTGGCCCCGCTGCGAGCAGAGATGAAGCTCTGCTCGCAGATCTGTGCGACCC 2898  
3093 AGTGTGCAATAGCTCAACCGCTCTGTTCGCGAAGCAACCGAAGAGTTCCTCGAGCA 3152  
2899 -GTTGCCCAACGCTGGATCGCTCTCTTCCGGAAGCTGCGGAGAAATTGCGAGAGCA 2957  
3153 CGTGTGCGCTTCGCGACACCTCTGCGCTGATGATCGTGAATCTTCTAGCGGCTGCT 3212  
2958 CGCGCTCAGTTTGGCGATACCAACAACTCGGTGACGAGAAATTCCTCTATGGCTCAA 3017  
3213 CGAAGCGCCGAGACTTTGATCCGCTCGCAGATGTGCGCACCCCACTGCTTGTTCGCT 3272  
3018 AGAAGCCAAAGAAACGCTCATTCGTACGCAACAGTAGCTAGCTGCCAATGCTGCTGCT 3077  
3273 GATGGATCTCTGAGCAGAGATTAAGGATGCGCATGTTGTGGCAGCTCAAGCG 3332  
3078 CGATGCGCTCGCGAGCCGCGCAAGGCGATGCGCAACGTTGTGTGCAATGTCAACGG 3137  
3333 CCAGATCCCGCAATCGGTGTGCGTACCGCTCCGTTGAGTCTGTACCGCAACCGCAGA 3392  
3138 CCAGATTCGCCCACTCTCTGCTGGTATCGCAGCGTAGAGTCAGTGACCGCTCGTGA 3197  
3393 AAAGCAGATTTCTCCAAAGGCGCATGTTGCTGCAACCATTCGCTGCTGTTGTACCGT 3452  
3198 AAAGGCTGATCTCTTCCAGCGCGGACACGTCAGCAGCAACCATTCGCGAGGCTTGTCAAGT 3257  
3453 GACTGTTGCTGAAGGTGATGAGTCAAGGCTCGAGATGAGTCCGCAATCATCGAGGCTAT 3512  
3258 GACATTCGAGGAGGCGCTACCGTCAAGCGCGGATCCAGTTGCTGTCTATGAGGCAT 3317  
3513 GAAGATGGAAGCAACATCACTGCTTCTTGTGAGCGCAAAATCGATCGCGTGTGTTGCTCC 3572  
3318 GAAGATGGAAGCAACATCTCTGCCACACAGACGGAACCGTCCGACCGTATCGTCTTGAC 3377  
3573 TGCTGCAACGAAGGTGGAAGGTGGGACTTGTGATCGTCT 3611  
3378 GCGAGCCACCAAGTCAAGGTGGAAGCTTGTCTCTCTCT 3416

## RESULT 7

US-10-282-122A-25766  
; Sequence 25766, Application US/10282122A  
; Publication No. US20040029129A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: EPI/FA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25766  
; LENGTH: 3381  
; TYPE: DNA  
; ORGANISM: Mycobacterium avium  
; US-10-282-122A-25766

Query Match 37.6%; Score 1361.6; DB 12; Length 3381;  
Best Local Similarity 63.2%; Pred. No. 0;  
Matches 2131; Conservative 0; Mismatches 1234; Indels 9; Gaps 2;  
  
QY 238 AAGATCTTGTAGCAACCGCGGCGAATCGCGTCCGTCTTCCGTGCGAGCACTCGAA 297  
DB 10 AAGTCTGTGCGCAACCGCGGCGATCGCATCGCGCGTTCGCGGCGCTACGAG 69  
  
QY 298 ACCGGTGCAGCAGGTATTTACCCCGGTGAAGATCGGGGATCATTCACCGCTCT 357  
DB 70 CTGGAATGCCACCGTGGCGGTGTATCCCTACGAGGACCGCAATTCGGTGCACCGGTTG 129  
  
QY 358 TTGCTTCTGAAGTGTCCGATTTGACGGAAGGCTCACAGTCAAGCGTACCTGGAC 417  
DB 130 AAGCGCAGAGTCTTACCAAGATCGGTGAGAGGCGCACCCAGTCCGCGCTACTGTCC 189  
  
QY 418 ATCGATGAATTTATCGGTGCGAGCTTAAAGTTAAAGCAGATCCCATTTACCGGGATAC 477  
DB 190 GTGACGAGATCGTGGCAGCGGCGTGGCTGGCGGCGGATCGCATCTACCCGCGTAC 249  
  
QY 478 GGCTTCTGTGAATGCGCAGCTTGGCGGAGTGTGCGGAACCGGATTAATTTT 537  
DB 250 GGCTTCTGTGAGAAACCCGATCTGGCGGCGGTGGCGGCGCGGCGATCAGTTTC 309  
  
QY 538 ATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATTAAGTCTCGCGGCTAACCGCC 597  
DB 310 GTGGCCCCAGCGCGAGGTGCTCGAGTCAACCGGCGACAAGTCTCGGGCCATCGGGCG 369  
  
QY 598 GCGAAGAGCGTGTCTGCCAGTTTGGCGGAATCCACCCGAGCAAAACATCGATGAG 657  
DB 370 GCGCGGCCCGGGTCTTCCGGTGTCTGGCTCTCCGCGCGGCTCGACCTCGGTGCGAGAG 429  
  
QY 658 ATCGTTAAAGCGCTGAAGCGAGACTTACCCATCTTTGTGAAGCGAGTTGCGCGGTGT 717  
DB 430 TTGCTGAGCGCCGCGAGACGATGACGTTCCCGCTGTTCGTCAAGGGTTCGCGCGCGC 489  
  
QY 718 GCGGAGCGCGGTATGCGTTTGTCTTACCTGATGAGTTCGCAATTTAGCAAGAA 777  
DB 490 GCGCGCCGCGGAGTGCSCCGGTCAACGATCCGCGGCGCTGCGCGAGCGATCGAGGCG 549  
  
QY 778 GCATCTGTGAAGTGAAGCGGCTTTCGCGGATGGCGCGGTATATGTGCAACGCTGTGTG 837  
DB 550 GCCAGCGGTGAGGCGGAGTCCGCGTTTCGGGGAGCGCTCGGTGTTCTCGAGCAGGCGGTG 609

QY 838 ATTAAACCTCAGCATATTGAAGTGCAGATCCTTGGCGATCAGACTGGAGAGTTGTACAC 897  
DB 610 ATCAATCCGGGCACATCGAGGTCCAGATCCTGGCCGACACGACCGCAACCGTGTACAC 669  
  
QY 898 CTTTATGAAGTGTACTCTCAGTGCAGCGTGTGTCACCAAAAGTTGTGAAATTTGCGCCA 957  
DB 670 CTCTACGAGCGGAGTGTGAGGTGCGCGGCGCACAGAGGTGATCGAGATCGCCCG 729  
  
QY 958 GCACAGATTTGGATTCAGAACTGCGGTGATCGATTTTTCGGGATGAGTAAGATTTTCG 1017  
DB 730 GCGCCCAACCTGTGATCGCGCTGCGCAACGATCTGCGCCGACGCGGTGGGCTTCGCA 789  
  
QY 1018 CGCTCATTGTTTACAGGCGCGGAAACCGTGAATTTCTTGGTGCATGAAAGGCGCAAC 1077  
DB 790 CGAGCATCGGTACACCTGGCGGCGACTGTGAGTTCTGCTGAGAAACGCGGCAAC 849  
  
QY 1078 CAGCTTCTCATCGAAATGAACCAAGTATCCAGTTTGAAGCAACCGGTGATGAAAGTTC 1137  
DB 850 CACGTGTTTCAGAGATGAACCGCGCATCCAGTGTGAGCAACAGGTGACCGAGGATC 909  
  
QY 1138 ACCGAGTGCACCTGTGTGAAGGCGCAGATCGCTTGGCTGCTGTGCAACTCTTGAAGGAA 1197  
DB 910 ACCGAGTGCACCTGTGTGTCGCGCAGCTGCGCATCGCTTCCGCGCAGAGCTGGAAGAG 969  
  
QY 1198 TTGGTCTGACCCCAAGATAAGATCAAGACCGAGTGCAGACTGAGTGCAGATCAC 1257  
DB 970 ATTGGCTTGAGCCAGGATTCGGTCACTCAAGCGCGCGCGCTGCACTGCGGATCAC 1029  
  
QY 1258 ACAGAGATCCAAACAGCGTTCGCGCCAGATACCGGAATATCACCGGTACCGCTCA 1317  
DB 1030 ACCGAGACCCCGCAACGGTTTCGCGCGGACACCGCGCGCATCACCGCTACCGCAC 1089  
  
QY 1318 CCAGCGAGTGGCGTTCGTCTTACGCTGCACTCAGTCTGGTGGCGAAATCACCGCA 1377  
DB 1090 CCGGTGGCGGCGCATCCGCTGCGACCGCGGACCAACCGCTGGCGCGGAGATCAGCGCG 1149  
  
QY 1378 CACTTTCATCTCATGCTGTGAAATGACCTGCGGTGCTTCCGACTTTGAACTGCTGTT 1437  
DB 1150 CACTTCGATTCGATGCTGATCAAGTCACTGCGCGGAGCGGACTTCCCCACCGCGGTG 1209  
  
QY 1438 GCTGTGCACAGCGCGGTTCGCTGAGTTTCACTGCTGCTGTGTTGCAACCAATTTGGT 1497  
DB 1210 CGGCGGCGCGCGCGGTTCGCGCAATTCGCACTCGCGCGCTCTCGAGATATCCCG 1269  
  
QY 1498 TTCTTGGCTGCTTTCGCGGGAAGAGATCTTCACTTCCAGCGCATCGCCACCGGATTC 1557  
DB 1270 TTCTTGAAGCGGTCTCTGACGACCCGAGTTTTCAGGCTGGGCGCATCACACCTGTT 1329  
  
QY 1558 ATTGCGCATCACCGCACCTCTTTCAGGCTCCACTGCTGATGATGAGCAGGAGCGCATC 1617  
DB 1330 ATCGAAGCGCGCGCGGTTCCTCACCGCGCGAGTTGCGCCGACCGCGGACCCAGATC 1389  
  
QY 1618 CTGATTAATTGGCAGATGTCCCGTGAAACAGCGCTCATGTGTGCTTCCAAAGATGTT 1677  
DB 1390 CTCAACTACCTTGGCGGAGTGCAGCGTCAACAGCGCGACGCGCGGCGCGCTCGCGGTG 1449  
  
QY 1678 GCAGTCTCTATCGATTAAGTGCCTTAAATCAAGATCTGCGACTGCGCACGCGGTTCGGT 1737  
DB 1450 TACCGCATGACAGTGTGCGCATGCACTGCGCGGAGCGCGCGCGGCTTCCAG 1509  
  
QY 1738 GACCGCTGAAGCAGCTTGGCCAGCGCGTTCGCTCGTGTATCTTCGTGAGCAGGACGCA 1797  
DB 1510 CAGCGTTTGAACCGAATCGCGCGGAGCGATTCGCGGCTTGGCTGCGCGAGTGGCGGCG 1569  
  
QY 1798 CTGGCAGTTACTGATACCACTTCGCGGATGCAACCACTTCTTGTTCGACCCGAGTC 1857  
DB 1570 GTGCGGTGTCAGCAGACCACTTCGCGGAGCGGACCAAGTCTGCTGCGGACCGCGGTG 1629  
  
QY 1858 CGCTCATTCGACATGAAGCTTCGCGGAGAGCGCTGCAAAAGTGTGCTCTGAGCTTTTG 1917  
DB 1630 GCGACAGCGGCTGCTCAAGTGTGCGACCTTACATCGACGAGCATTCGCGGAGCTGTTG 1689

1918 TCCTGGAGGCTTGGGGCGGCGGACCTACGATGTGGGATGCGTTTCCTCTTTGAGAT 1977  
1690 TCGGTGGAGTGTGGGGCGGAGCGCACTTACGACGTGGCGTTCGCGTTTCTCAAGAGGAC 1749  
1978 CCGTGGGACAGGCTCGACGAGCTGGCGGAGCGATGCCGAATGTAAACATTCAGATCTG 2037  
1750 CCGTGGGAGCGCTGACGCGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATCTG 1809  
2038 CTTTCGGGCGGCAACACCGTGGGATACACCGGTACCCAGACTCGGTCTGGCGGCGTTT 2097  
1810 CTGCTGGGCGTAAACCGCTGGGGTACACGCGCTTATCCGAGACGGTCAACGCGCTTC 1869  
2098 GTTAAAGAGCTGCCAGTCCCGGCTGGACATCTTCCGCATCTTCGACGCGCTTAAAGAC 2157  
1870 GTCCGGAGCGACCGAGACCGGATCGACATCTTCGCACTTCGAGCGCTTGAACAC 1929  
2158 GTCTCCGAGATGCTTCAGAAATCAGCGAGTCTTGGAGACACACCGCGGTAGCGAG 2217  
1930 GTCCGATCGATCGCCCGGACATCGACGCGTTCGCGAAACCGGAAGCGCATAGCCGAA 1989  
2218 GTGGCTATGGCTTATCTGGTGTATCTCTGATCCAAATGAAAGCTCTACACCTGGAT 2277  
1990 GTGGCGATGTCTACACCGGACCTGTTCGATCCGCGGAAGAGCTTACACGTGGAC 2049  
2278 TACTACCTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGCTCACATTTTGGCCATTAA 2337  
2050 TATTACCTGAAGCTGGCGGACAGATCGTCAAGCGGGCGGCGACGCTGTGGCGATCAAG 2109  
2338 GATATGGCTGTCTGCTTCGCGGACGCTGGGTAAACAGCTGGTCAACGCGACTGCGCGT 2397  
2110 GACATGGCGGCTGCTGGCGGACCGCGGCGGCGAGCTGGTGTGGCGCTGAAGTCG 2169  
2398 GAATTCGATCTCCAGTGCACGTGCACACCCAGACACTGGCGGTGGCGAGCTGGCAACC 2457  
2170 CGGTTCCGACTCGCGGTGCACGTGCACACCCAGACACTGGCGGCGGCGAGCTGGCGACC 2229  
2458 TACTTTGTGAGCTCAAGCTGTGCAGATGTGTGACGTGTTCGCGACCACTGTCT 2517  
2230 TAGTGGCGGCTGGCAGCGCGGCGGAGTGCCTGCGAGCGCGCGCGCGCGCTGGCC 2289  
2518 GGCACCACTCCCGACCACTCTGTCTGCAATTTGTGTGCAATTCGCGCAACACCGTGC 2577  
2290 GGAAACCAACGACGCGGCACTATCGTCTGATCGTGGCGCGCGCGGCGCAACACCGAGTAT 2349  
2578 GATACCGTGTGAGCTCGAGCTGTCTGACCTCGAGCGCTACTGGAGGAGTGGCG 2637  
2350 GACACCGGTGTGCTCTCGCGGCTGTGCTGATCTGGAGCGCTACTGGAGGCGCTGCGA 2409  
2638 GGAAGTACGCTCCCTTCGAAATCGGCTCTTCGCGCGGCGAGCGGCGGCTGTATCACAC 2469  
2698 GAATCCGAGGCGGACAGTGTTCACCTGCGTGCACAGGCGACCGCACTGGCGCTTGG 2757  
2470 GAGATCCCGGCGGCGCAATTTGCGAATCTGCGTTCAGAGCGGAGTTCGCGTGGGATCGCG 2529  
2758 GATCGTTTGGAACTCATCGAAGACAACTACGCGAGCGGTTTAAATGAGATGTGGAGCGCCA 2817  
2530 GACCGGTTCCGAGACATCGAAAGCGCTACGCGGCGCGGCGAGCGGATTTTGGGACACCTG 2589  
2818 ACCAAGTACCCCATCTTCAGGTTGTGGGACCTCGCACTCGACTCGTTGGTGG 2877  
2590 GTCAAGGTCAACCGTCCAGCAAGTGTGGGCGACCTGGCGCTGGCGCTGGTGGGTGG 2649  
2878 GGTGTGGATCCAGCAGACTTTGTGCGGATCCACAAAGTAGACATCCAGACTCTGTC 2937  
2650 GGTGTGGGCGGCGAGACTTCGCGGAAGACCCCTCCGCTACGACATCCCGGATTCGGTG 2709  
2938 ATCGCGTTCTCGCGGCGAGCTTGTGTAAACCTTCAGGTGGTGGCGAGCACTGGCG 2997  
2710 ATCGGCTTTCTGCGCGGAGCTCGGTGACCCCGCGGCGGCTGGCGAGCGGTTCGCG 2769  
2998 ACCCGGCGACTGGAAGCGCGCTCCGAAGCAAGGCACTCTCGACGGAAGTTCTCGAGGAA 3057

2770 ACCAAGCTTACAGGCGCGGGCGGCGCAAGCGGAGCAGCGCTGACGCGCGAGGAC 2829  
3058 GAGCAGGCGCACTCGAGCGCTGATTCGAAGAACTTCGAATAGCTTCAACCGCTG 3117  
2830 GAAGCGGCTTGGCGCGCGCGGCTG-----CGCGACGCGAGGCGCTGAACCGGCTG 2883  
3118 CTGTTTCCCAAGCAACGAGAGTTCCTCGAGCAGCGTTCGCGGCTTCGGCAACACCTCT 3177  
2884 CTGTTTCCCGGACCGACAGAGGCTCGAGGACACCGCGAGCAGTACGGTGACACTCG 2943  
3178 GCGCTGATGATCGTGAATTTCTTACCGCTGTGTCGAAGCGCGGAGACTTTGATTCGCG 3237  
2944 GGGCTGAGCGCAACCGAGTTCTTACGATTCGCGCAGGCGAGCAACACCGCTCGAA 3003  
3238 CTGCCAGATGTGGCAGCCCACTGCTTTCCTGCTGGATGCGATCTCTGAGCCAGCAGT 3297  
3004 CT---CGAGCGGCGCTGGAGTTGCTGATCGGCTTGGAGGCCATCTCGATCCGACGAA 3060  
3298 AAGGATATCGCAATTTTGTGCGCAACGTCACCGCCAGATCCGCCCAATGCGTGTGCT 3357  
3061 CGCGGATCGCAGCGTGTGATGTCATTTCTCAACGCGCAGCTCGCGCGGCTGGTGGTGG 3120  
3358 GACCGCTCGTTGAGTCTGTACCGCAACCGCAGAAAGGCGAGATTCCTCCAACAAGGCG 3417  
3121 GACCGGAGATCGCCATCGACGTGCGCGCGGCGAGAGGCGCGCGCGCTTACCAG 3180  
3418 CATGTTGCTGCACCAATTCGCTGTGTGTCACCGTACTTTGCTGAAAGGTGATGAGTTC 3477  
3181 CACATCGTCCCGCTTTCGCGGCTGTGACGCTCGCGCGGAGTTCGCGCGGCGAGTTC 3240  
3478 AAGCTGGAGTGCAGTTCGCAATCATCGAGGCTATGAGATGGAAGCAACATCACTGCT 3537  
3241 GAGCGCGCGCAGCACTCGGAGCAGTTCGAGGCGGATGAGATGGAAGCGCGCTTACCAG 3300  
3538 TCTGTTGACGCGCAAAATCGATCGCGTGTGTTGCTTCCTGTCGAACGAGTGGAAAGTGGC 3597  
3301 CGAAGTTCGCGCAAGTGGCGGCAATCGCGCTGTCGCGACCGCTCAATCGAGGCGCGC 3360  
3598 GACTTATGCTGT 3611  
3361 GACCTCTGATGCT 3374

## RESULT 8

US-10-282-122A-28626

; Sequence 28626, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITEA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

```

; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28626
; LENGTH: 3384
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28626

Query Match      35.7%; Score 1294.4; DB 12; Length 3384;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 2099; Conservative 0; Mismatches 1276; Indels 9; Gaps 3;

QY 238 AAGATCTTGGTAGCAAAACCGCGCGAAATCGCGTCCGTCGTTTCCGTCGAGCACTCGAA 297
DB 10 AAGGTCTCGTCGCAATCGCGGAGATCGGATCCGGGCTTTCGTGCGGCTACGAA 69
QY 298 ACCGTCGAGCAGTAGTATTTACCCCGTGAAGATCGGGATCAATTCACCGCTCT 357
DB 70 CTGGCGTCGGAACCGTGGCGTTATCCGTACGAGGACCGCAATTCGACGACCGTCTC 129
QY 358 TTTGCTTCTGAAGTGTCCGATTTGTACCAAGGCTCACAGTCAAGGCGTACCTGGAC 417
DB 130 AAGGGGACGAGTCTTACCAGATCGGCGACATCGGTCAACCGGTGCATATACCTGTGC 189
QY 418 ATCGATGAATATTCGTCGAGCTAAAGTTAAAGCAGATGCGATTAACCGGGATAC 477
DB 190 GTCGACGAGATCGTCGCGACCGCCGTCGCGGGGTCCGACGCTATCTACCTCGGTAC 249
QY 478 GGTTCCTGTCTGAATATGCCAGTTCGCGGAGTGTGCGGAAACGGCATTAATCTTT 537
DB 250 GGGTTCTATCGGAGATCCGATCTGGCTGCGCATCGCGCGGCGGCGCATCGCTTC 309
QY 538 ATTGGCCCAACCCAGAGTTTGTGATCTCAACGGTGAAGTCTCGCGGGTAAACGCC 597
DB 310 GTCGTCCTCCAGCGCGAGTGTCTGAGCTGGCTGGGAAATAGTCTCGCGCATCGCGGCG 369
QY 598 GCGAAGAGGCTGTCTGCCAGTTTTCGCGAATCCACCCCGAGCAAAACATCGATGAG 657
DB 370 GCCCGGAGCCGTTGCTCCGCTGATGTCCTCGCGCCGCTCGGCTCGGTGCGAGAA 429
QY 658 ATCGTTAAAGCGTGAAGCCAGACTTACCCATCTTTGTGAAGGCAAGTTAGCAACAGAA 717
DB 430 CTGCTGTGCGTTCGCGCGCATGCCGTTTCCGTTTCTGTCGTCGAGGCAAGTTGCGGTGC 489
QY 718 GCGGACCGCTATGCTTTGTTCTTACCTGATGAGTTCGCAATTAGCAACAGAA 777
DB 490 GGGGCCCGGGGTATGCTGCTGTGCGGCAATATCGCGCGCTTCGAGGCGGATCGAAGCC 549
QY 778 GCATCTCGTGAAGTGAAGCGGCTTTCGCGATGGCGGATATATGTCGAAACGTGCTGTG 837
DB 550 GCCAGCCGGGAAGCCGAGTCCGCTTCGCGGACCCGACGCTCTATCTCGAGCAGGCAATG 609
QY 838 ATTAACCTCAGCATATTAAGATGAGATCTTTCGCGATCACTGAGAGATGTTACAC 897
DB 610 ATCAATCACGCCCATCGAGTGCAGATTCCTGGCGGACCACTTCGCGGACGTGATCCAT 669
QY 898 CTTTATGAACGTGACTCTCACTGACGCTGTCAACCAAAAGTTGTCGAAATTCGCCCA 957
DB 670 CTCTATGAGCGTACTGAGTGTGACGCTGCCATCAGAGTCTACGAGCTGGCGCC 729

```

```

958 GCACAGCATTTGGATCCAGAACTCGGTGATCGCATTTGTGCGGATGCAATGTTCTTGC 1017
DB 730 GCGCGCACCTGGAGCGCGAGTTGCGTTTACAGATGTGCGTCGATGCGGTTCGCTTGC 789
QY 1018 CGCTCCATTGGTTACAGCGCGCGGAAACCGTGAATTTCTTGGTCGATGAAAGGCAAC 1077
DB 790 CGCCATATCGGGTACAGCTGCGCGGACCGTCTGAGTTCTGCTGAGCAGCGAGGAG 849
QY 1078 CACGTCCTTCATCGAAATGAACCCAGATATCAGGTTGAGCAACCGTGTGAGAGTGC 1137
DB 850 TATGTCCTTCATCGAGATGAATCCGCGGTTTCAAGTGGAGCACACGCTGACCGAGATT 909
QY 1138 ACCGAGGTGGACCTGGTGAAGCGCGAGATGCGTGGCTGTGCTGCTGCAACCTTGAAGAA 1197
DB 910 ACCGAGCTGACCTGGTCCGACGACCTGGCATTTGCCGCGGGGAGACGCTCGAACAA 969
QY 1198 TTGGGTCTGACCCCAAGATGAATCAAGACCCACCGTGTGAGCACTGCGATCACC 1257
DB 970 TTGGGCTTCGCGCAGGAGACATCGCACCGCATGCGCGCACTACAGTGCAGTCAACC 1029
QY 1258 ACGGAAGATCCAAACCAACCGGTTCCGCCACATACCGGAATATCACCGCTACCGTCA 1317
DB 1030 ACGGAGATTCGCCACCGGTTTCGCCCGGACACGCGCGGATCAGCGCTTGGGAC 1089
QY 1318 CAGCGGAGTGGCGTTCGTTTGAAGTGCAGTCTGAGTTCGAGTTCGAGTTCGAGTTC 1377
DB 1090 GCGCGCGTGGCGGTTCGCTGAGCGGACCAACCTGGCGGCAAGAAATCAGCCG 1149
QY 1378 CACTTGTGATCTGCTGTAATGACCTGCGGTGCTTCCGACTTTCGAACTGCTGTT 1437
DB 1150 TACTTCGACTTCATGCTGCTCAAGTGAACCTGTCGCGGCGGTGACCTCCCTACCGAGTG 1209
QY 1438 GCTCTGTCACAGCGCGCTTGGCTGAGTTTCAACCGTGTCTGCTGTGCAACCAATTCGT 1497
DB 1210 AGCGTGTGCGCGCGGATCGCGAGTTCCGATCCGCGGATATCGACGAATATTCG 1269
QY 1498 TTCTTGGTGGTTCGTCGCGGAGAGGACCTTCAATTCAGAGCATCGCCACCGATTC 1557
DB 1270 TTCTGCAAGCGGTCTCGATGACCCGACTTCGAGCGGCGCGGTTCACCACTGCTTC 1329
QY 1558 ATTGCGCATCACCCGACCTCTTCAGGCTCCACTGCTGATGATGAGCAGGAGCGATC 1617
DB 1330 ATTGATGAGCGCGCGAGTCTGACCGCGCGCTCGCGGACCGCGGCGACCAAGATC 1389
QY 1618 CTGGAATTTGCGAGATGTCACTGGAAGAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTG 1677
DB 1390 CTAACTTCTTGGCGGATGTCACTGCAACCAACCGTATGCTGCTGCTGCTGCTGCTGCTG 1449
QY 1678 GCAGCTCCTATCGATAAGCTGCTAAACATCAAGGATCTGCCACTGCCACGCGGTTCCGCT 1737
DB 1450 TACCGGAGCAAGCTGCCGATCTTGTATCTGCGGCGCGCACCAACCGCGGCTCAG 1509
QY 1738 GACCGCTGAAGAGCTTGGCGGAGCGGCTTGTGCTGCTGATCTCCGCTGAGAGAGAG 1797
DB 1510 CAGCGACTAGTCAAGTTGGGCGCGGAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1569
QY 1798 CTGCGAGTTACTGATACCACTTCGCGCATGACACCACTGCTTGTGCTGCTGCTGCTGCTG 1857
DB 1570 GTGCGGTCACGATACCACTTCGCGGATGCTTCCGAGTGTCTGCTGCTGCTGCTGCTGCTG 1629
QY 1858 CGCTCATTTGCACTGAAGCTTGGCGCAGGCGCTGCAAGAGTGTGCAAGTGTGCTGCTGCTGCTG 1917
DB 1630 CGCACCGAGCTGTGCGGCTGCGACCGTATCTCGCGCGGACCATGCGCGAGCTGTTG 1689
QY 1918 TCCGTGAGGCTTGGCGGCGGCTTACGATGTCGAGTGTGCGGATGCTGCTTCTTTCGAGAT 1977
DB 1690 TCCGTGAGGCTTGGCGGCGGCTTACGATGTCGAGTGTGCGGCTGCTGCTTCTCAAGAGAT 1749
QY 1978 CCGTGGGACAGGCTCGAGAGCTGCGGAGGCGGATGCGGAATGTAAACATTCAGATGCTG 2037
DB 1750 CCGTGGGACAGGCTGCGGACATGCGGTGCGCAATGCCAATATCTGCTTGCAGATGCTA 1809
QY 2038 CTTGCGGCGCGCAACACAGTGGGATACACCCGCTACCCAGACTCCGCTGCTGCGCGGCTT 2097

```

1810	TTTCGGGGCGCGCAATACCGTGGGGCTACACGCGGTACCCGGAAATCGTGAAGCTCGCGCGTTT	1869
2098	GTTTAAGAAAGCTGCCAGCTCCGGCGGTGGACATCTTCGGCATCTTCGACGCGCTTAAACGAC	2157
1870	GTCAGAAGACCAACGCCACCGGTATCGACATCTTTCGTATCTTCGACGCGCTCAACAAC	1929
2158	GTCTCCAGATGGGTCCAGATTCGACGAGTCTCTGGAGACCAACACACCGCGGTAGCCGAG	2217
1930	ATCGAGTCGATGGCTCGCGGATTCGACGAGTACGCGAAACAGGTTCTCGATAGCAGAA	1989
2218	GTGGCTATGGCTTATTTCTGGTGAATCTCTGTATCCAAATGAAAAAGCTCTACACCCCTGGAT	2277
1990	GTCCGATGTGCTACACGGGCGACCTTACCGATCCGGGTGAACAGCTGTACACGCTGGAC	2049
2278	TACTACTTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAAG	2337
2050	TACTACTGTAACCTGGCTGAGCAGATCTGGAGCGCGCGGCCCATGTGCTGGCGATCAAG	2109
2338	GATATGGCTGGTCTGCTTCGCCCAAGCTGGGGTAACCAAGCTGGTCAACCGCATCTGGCGCGT	2397
2110	GATATGGCTGGAATGTTGGCGCGCGCGCGCTCAAAAGTTGGTCAAGCGCCCTGGCGCACT	2169
2398	GAATTCGATCGCATGTCAGTGGACGTCACCCACGACACTGCGGGTGGCGCAGCTGGCAACC	2457
2170	CGCTTGGACCTGCCCGTTCACTTCGACACCCACGACACCGGGTGGCGCAGCTCGGCAGC	2229
2458	TACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGAAGGTGCTTCCGACCACTGTCT	2517
2230	TATGTGGCGCTGTGGCACCGCGCGGCGATGCCGTGACGCGCGCGCGCGCTGGCG	2289
2518	GGCAGCACCTCCAGACCATCCCTGTCTGCATTTGCTGTCATTTCCGCGCACACCGTCGC	2577
2290	GGAAAGCAGCAGACCGCGCTGAGCTCGATCGTTGCTGCCCTGCCACACCGAGTAC	2349
2578	GATACCGGTTGAGCTTCGAGGCTGTTTCTGAACCTCGAGCGCTACTGGGAAGCAGTGGCG	2637
2350	GACACCGCCCTGCTCTTCGGCGGTGCGGCCCTGGAGCCGTACTGGGAGGCGTTACGA	2409
2638	GGAGTGTACCTGCATTTGAGTCTGGAAACCCAGGCCCAACCGGTCCGCTTACCGCCAC	2697
2410	AAGTGTATCGCGCTTCGAGTCTGGGTTCCGGGGCCGACGGGGCGGTTTATCACCAC	2469
2698	GAATCCAGCGCGACAGTTGTGCAACCTCGGTGCAACAGGCCACCGCATCTGGGCCCTTGGG	2757
2470	GAGATTCGGGCGGCCAACTGTCCAAATCTGCGCCAGCAAGCAATGCTCTTGCTTGGGA	2529
2758	GATCGTTTCGAATCTACGAAGACAACTACGACGCGTTAATGAGATGCTGGACACGCCCA	2817
2530	GATCGATTCGAAGATTCGAAGGCCCTACGCGGCGCGCACCGAGTGTGGGCGAGCTG	2589
2818	ACCAAGTCAACCATCTCTCAAGGTTGTGGGCACTTCGCACTCCACCTCGTTGGTGGCG	2877
2590	GTTTAAAGTTCACGCCAAACGTCGAAGTGGTGGCGATCTGGCGCTGGCACTGTGTCGGCC	2649
2878	GGTGTGATCCAGCAGACTTGTCTGCCGATCCACAAAGTACAGACTCCCGAGACTCTGTC	2937
2650	GGTGTCACTGCAGACGAATTCGCTCCGATCCAGCGCGATTTGGCATCCCGGAATCGGTA	2709
2938	ATCGCGTTCTCGCGCGGAGCTTGGTTAAACCTTCAGGTGGCTGGCCAGAGCCTACCTGCGC	2997
2710	CTCGGATTTCTGGGGGCGAGCTGGGTGATCCGCGCGCGGTGGCCGAACCGCTGCGC	2769
2998	ACCGCGCACTGGAAGCCGCTCCGAAGGAAGGCACTCTGACGGAAGTTCTCTGAGGAA	3057
2770	ACTCGGCGCTGGCGGTTCGGGGCGGCCAGGCCACT--GCGCAATTTGGCGGCGGAC	2826
3058	GAGCAGGGGCACCTCGACGCTGATGATTCGAAGGAAGTTCGCAATAGCCTCAACCGCTG	3117
2827	GACGAGATTGCCCCCTATCGTGGTCCGAGGCCAAG--GTCAGGCCACCCCTGAACAGGCTG	2883
3118	CTGTTCCCGAAGCCAAACGAAGAGTTCTTCGAGCACCGTTCGCCCTTCGGCAACACCTCT	3177

## RESULT 9

RESULTS  
IIS-10-282-122A-26500

US-10-282-122A-26500  
 . Sequence 26500 Application US/10282122A

; Sequence 26500, Application US/

; Publication No. US20

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl

APPLICANT: Ollibell, Ralf  
APPLICANT: Zygkind, Jud

APPLICANT: Wall Daniel

APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

**; APPLICANT: Yamamoto**

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

**TITLE OF INVENTION: Identification of Esse**

FILE REFERENCE: ELITRA-034A

; FILE REFERENCE: ELLIRA.034A  
: CURRENT APPLICATION NUMBER: IIS/10/28

```

; CURRENT APPLICATION NUMBER:
; CURRENT BILLING DATE: 2003-0

```

; CURRENT FILING DATE: 2003-02-20  
 : PRIOR APPLICATION NUMBER: 504191 078

;  
PRIOR APPLICATION NUMBER: 60  
PRIOR FILING DATE: 2000-03-23

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60

; PRIOR APPLICATION NUMBER: 60/230,333  
: PRIOR FILING DATE: 2000-09-06

```

; PRIOR FILING DATE: 2000-09-0
; PRIOR APPLICATION NUMBER: 60

```

; PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000 08 08

; PRIOR FILING DATE: 2000-09-0

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-2

;  
; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO: 26500  
 ; LENGTH: 3381  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium bovis  
 US-10-282-122A-26500

Query Match 35.7%; Score 1292.2; DB 12; Length 3381;  
 Best Local Similarity 62.0%; Pred. No. 0;  
 Matches 2095; Conservative 0; Mismatches 1273; Indels 9; Gaps 3;

Qy	238	AGATCTTGTAGCAACACCGCGGAAATCGCGTCCGTTCCGTCAGAGCTCGAA	297
Db	10	AAGGTGCTCGCCAAATCGCGGAGATCGGATCCGGCCCTTTCTGTCGCGCTACGAA	69
Qy	298	ACGGTGCACCGAGGTAGTATTATCCCGCGGAAGATCGGGATCATTTCCACGCTCT	357
Db	70	CTGGGCTCGGAACCGTGGCCGTTTATCGGTACGAGACCGCAATTCGACGACCGTCTC	129
Qy	358	TTTGTCTTGAAGCTGTCCGCAATGGTACCGAAGGCTCACAGTCAAGCGTACCTGGAC	417
Db	130	AAGCGGACGAGCTTACCAAGATCGGCGACATCGGTCAACCGGTGATCATACCTGTCG	189
Qy	418	ATCGATGAATATTCGTGACGATTAAGTTAAAGCAGATGCCATTTACCGGGATAC	477
Db	190	GTCAAGAGATCTGCGGACGCGGCCGTCGCGGGTGCAGACGCTATCTACCTCGCTAC	249
Qy	478	GGCTTCCTGTGAAATGCCCCAGTGTGCCCGGAGTGTGCGGAAACCGCATTTACTTTT	537
Db	250	GGGTTTCTATCGGAGATCCGATCTGGCTGCGCATGCGCGCGGCGATCAGCTTC	309
Qy	538	ATTGGCCCAACCCAGAGTTCTTGATCTCACCGGTGATAGTCTCGCGGGTAAACGCC	597
Db	310	GTCCGTCCAGCGCGAAGTGTGAGTGGCTGGGAATAGTCTCGCCCATTCGCGGCG	369
Qy	598	GGCAAGAAGCTGTGTGCGAGTTTGGCGGAATCCACCCCGAGCAAAACATCGATGAG	657
Db	370	GCCCGCGAGCGGCTTGCCTGCTGATCTCTCGCGCGCTCGGCTCGGTCAGAA	429
Qy	658	ATCGTTAAAGCGCTGAAGCGCAGATTACCCCATCTTTGTGAAGGAGTTGCGCGTGT	717
Db	430	CTGTCTCGTTGCGGCGCGCATGCCGTTTCCGTTGTTCTCAAGCAGTTGCGCGTGC	489
Qy	718	GGCGACGCGTATGCGTTTGTGTTTACCTGATGAGCTTCGCAATTTAGCAACAGAA	777
Db	490	GGGGCGGGGTATGCGTCTGTCGCGGATATCGCGCGCTTCGAGGCGATCGAAGCC	549
Qy	778	GCATCTCGTGAAGTGAAGCGGCTTTCGCGATGGCGCGTATATGTGCAACGTGTGTG	837
Db	550	GCCAGCGGGAAGCGAGTCCGCTTCGCGGACCCGACGCTCTATCTCGAGCAGGCA	609
Qy	838	ATTAACTCTAGCATATTGAAGTGCAGATCTTGGCGATCACACTGAGAGTTGTACAC	897
Db	610	ATCAATCAACCCACATCGAGTGCAGATCTTGGCGGACCACTCGCGGACGTGATCCAT	669
Qy	898	CTTTATGAAGCTGATGCTCATCTGACGCTGTGCACCAAAAGTTGTGAAATTCGCCCA	957
Db	670	CTCTATGAGCGTGACTGCAGTGTGACGCTGCCCATCAGAAGGTCTACGAGCTGGGCC	729
Qy	958	GCACGATTTGGATCCAGACTCGGTGATCGCATTTGTGGGATGAGTAAAGTTCTGC	1017
Db	730	GCGCGCACCTTGGACGCGGAGTTGCGTTACAGATGTCGTGATGCGGCTTCGCC	789
Qy	1018	CGCTCCATTGCTTACCGAGGCGCGGAAACCGTGGAAATTTTGTGTCGATGAAAGGGCAAC	1077
Db	790	CGCCATATCGGTACAGTGTGCGCGCACCGTCTGAGTTCTCTGTCGACGAGCGAGGGAG	849
Qy	1078	CACGTCTTCATCGAAATGAACCCAGGTATCCAGTTGAGCACCCGCTGACTGAAGATC	1137

Db	850	TATGTTCTTCATCGAGATGAATCCGCGGTTCAAGTGTGAGACACGTTACCGAGGATT	909
Qy	1138	ACCGAGTGCACCTGTGTGAAGCGCAGATCCGTTGCTGCTGCTGCAACTTGAAGAA	1197
Db	910	ACCGAGTGCACCTGTGTGCGCAGCAGCTCGCATTTCCCGCGGAGAGCGTTCGAACA	969
Qy	1198	TTGGTCTGACCCAGATAGATCAAGACCCAGCTGACACTGACGTGCGGATCAC	1257
Db	970	TTGGGCTGCGCGCAGGAGGACATCGCACCGCATGTGCGCGCATACAGTCCCGATCAC	1029
Qy	1258	ACGGAAGATCCAAACAAACGCTTCCGCCAGATACCGGAATAATACCGGCTACCGTCA	1317
Db	1030	ACCGAGATCCGGCAACGCTTCCCGCGGACACGCGCGGATCAGCGGTTGCGCAC	1089
Qy	1318	CCAGCGGAGTGGCGTTCGTTGACGTCGAGCTCAGCTCGGTGCGGAAATCACCGCA	1377
Db	1090	GCCGCGGTGCGGTGTCGCGCTGACGCGCAGCACMACTTGGCGCAGAAATCAGCCCG	1149
Qy	1378	CACTTTGACTCCATGCTGTGTGAAATGACCTGCGGTGTTCCGACTTTGAAATGCTGTT	1437
Db	1150	TACTTCGACTCCATGCTGTGTCAGTCACTGTGCGGCGGTGACCTCCCTACCGAGTG	1209
Qy	1438	GCTGTCGACAGCGCGGTTGGCTGAGTTCACTGTCGTGCTGTTGCAACCAATTTGT	1497
Db	1210	AGCGTGTGCGCGCGGCGATGCGGAGTTCCGATCCGCGGGGTATCGAGATATTTCCG	1269
Qy	1498	TTCTTGTGCTGTGCTGCGGAAAGAGACTTCACTTCCAAAGCGCATCGCCACCGGATTC	1557
Db	1270	TTCTTCAAGGGTCTTGGATGACCCGGACTTCGAGCGGCGCGGTCACCACTGCTTC	1329
Qy	1558	ATTGCGGATACCCGCACTCTCTTCAAGCTCCACTGCTGATGATGAGCAGGAGCGATC	1617
Db	1330	ATTGATGAGCGCGCGAGCTGCTGACCGCGCGGCTTCGCGCGGCGCACCGGACCAATC	1389
Qy	1618	CTGATTTACTTGGCAGATGTCAACGCTCAAGGCTCATGCTGTGCTGCTCAAAAGGATTT	1677
Db	1390	CTTAACTTCTTGGCGGATGTCAACGCTCAACAAACGCTATGCTGCTGCTCAACGATC	1449
Qy	1678	GCAGTCTCTATCGATAAGTGCCTAAACATCAAGATCTGCCACTGCGCACCGGTTCCG	1737
Db	1450	TACCGGACGACAAAGTGCCTGATCTGCGGCGCGCACCGCGCGGCTCCAAAG	1509
Qy	1738	GACGCTTGAAGCAGCTTGGCCAGCGGTTTCTGCTGATCTCCGTGACGAGAGCA	1797
Db	1510	CAGCGACTAGTCAAGTTGGGCGCGAAGATTTCTGTTGGCTGCGGAGTTCGCGCG	1569
Qy	1798	CTGGCAGTTACTGATACCACTTCCGCGATGACACCACTTTTCTTGGACCCGAGTC	1857
Db	1570	GTGCGGCTCACCGATACCACTTCGCGATGCTCAACAGTCTTACTGCTTACCGGATA	1629
Qy	1858	CGCTCATTTGCGAGCTTGGCGGAGCGCTGCGAAGCTGACTCTCTGAGCTTTG	1917
Db	1630	CGCACAGCGGACTGTGCGGGTGGCACCGTATCTTGGCGGACCATGCCGAGCTGTTG	1689
Qy	1918	TCCGTGAGGCTTGGGCGCGCGACCTTACGATGTCGATGCGGTTCTCTTTGAGAT	1977
Db	1690	TCCGTGAGTGTGGGCGGTGCGACTTACGATGTGCGCTGCGCTTCTCAAGAGGAT	1749
Qy	1978	CGGTGGGAGGCTCGACGAGTGGCGGAGGCTGCGGATGTAACATTCAGATGCTG	2037
Db	1750	CCCTGGGAACCGCTGGCCACACTGCGTGCAGCAATGCCCAATATCTGTTGAGATGTA	1809
Qy	2038	CTTCCGCGCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGCTGCGCGCGCTTT	2097
Db	1810	TTGCGGGCGCGCATACCGTGGCTTACACGCGCTACCGGAAATCGTGACGCTGCGGTT	1869
Qy	2098	GTTAAGGAGCTGCCAGTCCGCGGTGACATCTTCGCTATCTTCGACGCGCTTAAACGAC	2157
Db	1870	GTGCAAGAAGCAACAGCCCGGATCGACATCTTCTGATCTTCGACGCGCTCAACAC	1929
Qy	2158	GTCTCCAGATGCTGCCAGCAATCGACGCTCTTGGAGACCAACACCGCGGTAGCCGAG	2217

Db 1930 ATCGAGTCGATCGCGCGATCGACGAGTACGCGAAGACAGGTTCTGCGATAGCAGAA 1989  
Qy 2218 GFGGCTATGGCTTATTTCTGGTATCTCTCTGATCCAAATGAAGAGCTCTACACCTCGAT 2277  
Db 1990 GTGCGGATGTGTACACGCGGACCTTACCGATCGGGTGAACAGCTGTACACGCTGGAC 2049  
Qy 2278 TACTACCTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGCTCAATCTTGGCCATTAAG 2337  
Db 2050 TACTACCTAAAGATGGCAGAGATCGTGGAGCGGCGGCCATGTCTGGCGATCAG 2109  
Qy 2338 GATATGGTGTCTGCTTGGCCGAGCTGGGGTAACCAAGCTGGTCAACGCACTGGCGGT 2397  
Db 2110 GATATGGTGTGACTGTGCGCGCGCGCGCTCAAGGTTGGTCAAGCGCTTGGTCAAG 2169  
Qy 2398 GAATTCGATCTGCCAGTGCAGCTGCACACCCACGACACTGGCGGTGGCGAGCTGGCAACC 2457  
Db 2170 CGCTTGGACCTGCGGTTCACTCTGCACACCCACGACACACCGGTGGCGAGCTGGCCAGC 2229  
Qy 2458 TACTTTGTGAGCTCAAGCTGGTGAGATGTGTGTGAGCGGTCTTCCGCACTGTCT 2517  
Db 2230 TATGTGGCGCTTGGCAGCGCGGCGCGATGCGCTGCGACGCGCGCGCGCGCTGGCG 2289  
Qy 2518 GGCACACCTCCAGCACTCTGTCTGCAATTTGTGTGCAATTCGCGCACACCGCTGCG 2577  
Db 2290 GGAACGACGAGCGCGCGCTGAGCTGATGCTGTGCTGCGCTGCGCACACCGAGTAC 2349  
Qy 2578 GATACGGTTTGAAGCTCGAGCTGTTTCTGACCTCGAGCGGCTATGGGAAGCAGTGGC 2637  
Db 2350 GACACCGGCTGTGCTTTCGGCGGTGTGCGGCTCGAGCGGCTATGGGAGCGGTACGA 2409  
Qy 2638 GGACTGTACTCTCCATTTGAGTCTGGACCCAGCGCCCAACCGGTGCGGTCTACCGCCAC 2697  
Db 2410 AAGTGTATGCGCGGTTCAGTCTGGGTGCGCGGCGCGACGCGGCGGTATTACACAC 2469  
Qy 2698 GAAATCCAGCGGAGCAGTTGTCACCTGCGTGTGACAGGCGCACCGCACTGGGCGCTTGG 2757  
Db 2470 GAGATTCGCGGCGCACTGTCTATCTGCGCACAGCAATTTGCTTGTCTGGGA 2529  
Qy 2758 GATCGTTTGAATCTCAAGACACTAGCAGCGGTTAATGAGATGCTGGAGCGCCA 2817  
Db 2530 GATCGATTTGAAGATGAAGAGCGCTACGCGGCGCGCGACCGAGTGTGGGAGGCTG 2589  
Qy 2818 ACCAAGTCAACCCATCTCCAGGTTGTGGCGACTCGCACTCCACTCGCTTGGTGG 2877  
Db 2590 GTTAAGTCAAGCAAGCTTCAAGGTGTGCGGATCTGGCGTGGCACTGGTGGCGGC 2649  
Qy 2878 GGTGTGATCCAGCACTTGTGCGGATCCCAAAAGTACGACATCCAGACTCTCTC 2937  
Db 2650 GGTGTGATGACAGCAATTCGCTCCGATCCAGCGCGATTTGGCATCCCGGAATCGTA 2709  
Qy 2938 ATCGCGTTCTGCGCGCGAGCTTGTGTAACCTTCCAGGTGGGTGGCGAGCGCACTGGCG 2997  
Db 2710 CTGGAATTTCTGCGGCGAGCTGGGTGATCCGCGCGGGTGGCGCGAACCGGTGGCG 2769  
Qy 2998 ACCGCGCATGGAAGCGCGCTCGAAGCGAGCGACTCTGACGGAAGTTCTTGAGGA 3057  
Db 2770 ACTGCGCGCTGCGCGGTGCGCGGCGCGCGAGCCCACT---GCGCAATTTGGCGGCGAG 2826  
Qy 3058 GAGCAGCGCGCACTCGACGCTGATGATTTCAAGGAACGTGCAATAGCCTCAACGCGCTG 3117  
Db 2827 GACGAGATTCGCTTATCGTGGTGGAGGCGAAG---GCTCAGGCGCACCTGGAAGCGCTG 2883  
Qy 3118 CTGTTCCGAAGCGCAACCGAAGTTCTCTGAGCAGCGTGGCGCTTGGGCAACCTCT 3177  
Db 2884 TTATTTCCAGTCCCAACAAAGGAATTCATGAGCAGCGGGAAGCTTACGCGGAGCGTGG 2943  
Qy 3178 GCGCTGATGATCGTGAATTTCTTACGCGCTGTGTAAGCGCGGAGACTTTTGAATCCCG 3237  
Db 2944 CAATTGTGCGCAACCAAGTTCTTCTATGTTCTGGCCAGGTGAAGGATCGGTGAAG 3003  
Qy 3238 CTGCGAGATGCGCAACCCCACTGTTGTTGCGCTGATGCGATCTCTGAGCGAGAT 3297  
Db 3004 CTG---GAGCGTGGGGTGGAGCTGTTGATCGGGTGGAGGCCATTTCCGAACCCGAGCA 3060

Qy 3298 AAGGTTATCGCAATGTTGTGCGCAACGTCACGCGCCAGATCCGCGCAATCGTGTGGT 3357  
Db 3061 CGCGGATCGCAACGTTGATCTCTCAACGCGGAGCTCGCGCGGTCTAGTGGC 3120  
Qy 3358 GACCGCTCGTTGAGTGTCTACCGCAACCGCAGAAAAGGAGATCTCTCCACAGGCG 3417  
Db 3121 GACCGCAGCATTTGCCAGTGTCCGTCGCGCGCGGAGAAAGGCGGCAATCCCGGA 3180  
Qy 3418 CATGTTGCTGCAACCATTCGCTGTGTCACCGTGAATGTTGCTGTAAGGTTGATGAGTCT 3477  
Db 3181 CATGTCGCGCGCATTTGCGGAGTGTCTCAGGTTGGGTTGCTGCGGAGCGGTC 3240  
Qy 3478 AAGCTGAGATGAGTGTGCAATCATCGAGGCTATGAGATGGAAGCAACATCACTGCT 3537  
Db 3241 GCGCGCGGCAACCATCGCCACCATCGAGGCGATGAAGATGGAAGATCCCGGATCCCGC 3300  
Qy 3538 TCTGTTGACGCGCAAAATCGATCGCTGTGTTGTTCTGCTGCAACGAAAGTGGAAAGTGGC 3597  
Db 3301 CGGTTGCGGACACCGTGGAGCGGTTGGCGGTGTGCGACACCGCCAGGTTGGAGGCGGA 3360  
Qy 3598 GACTTGAATCGTGTGCT 3614  
Db 3361 GACCTGTGTTGTGTTGT 3377

## RESULT 10

US-09-917-800A-1566  
; Sequence 1566, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castle, Arthur  
; APPLICANT: Elashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917,800A  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,040  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,880  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: US 60/290,029  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/290,645  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US 60/292,336  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; PRIOR FILING DATE: 2001-07-09  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1566  
; LENGTH: 3945  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_012744  
US-09-917-800A-1566

Query Match 18.8%; Score 681; DB 9; Length 3945;  
Best Local Similarity 54.2%; Pred. No. 3.2e-157;  
Matches 1521; Conservative 0; Mismatches 1260; Indels 24; Gaps 6;



233 TCAAAAGATCTTGGTAGCAACCGCGCGAAATCGCGTCCGTGCTTTCCGTGACGAC 292  
144 TCAAGAAAGTAAATGGTGGCCACACAGAGGTGAGATGGCAATCCGAGTGTTCGTGGCTGCA 203  
293 TCGAAACGGGTGACGACCGTAGCTATTACCCCGTGAAGATCGGGATCATTCACAC 352  
204 CAGAGCTGGGTATCCGACAGTGGCTGTCTACTCGGAGAGGACACAGGCCAGATGCACC 263  
353 GCTCTTTTGTCTTGAAGCTGTCCGCAATGGTATCCGAAGGCTCACAGTCAAGCGCTACC 412  
264 GGCAGAAAGCTGATGAAGCTTACCTTATGGCGGTGGCTCCCTGTGCAAGCTACC 323  
413 TGGCATCGATGAATATCGGTGACGCTAAAGATTAAGCAGATGCCATTATCCCG 472  
324 TGCACATCCAGACATCATTAAGGTGGCCCAAGGAGATGTGTAGATGCTGTGCACCCCTG 383  
473 GATACGGCTTCCTGTCTGAAATGCCAGCTTGGCCCGAGTGTGGGAAACCGGCATTA 532  
384 GCTATGGTTCCTCTCAGAGAGACGACTTTGCCCCAGGCTGCCAAGATGCTGGAGTCC 443  
533 CTTTATTTGGCCAAACCCAGAGTTCCTGATCTCACCGGTGATAGTCTCGGCGGGTAA 592  
444 GATTCAATGGTCCAAAGCCAGAGTGTCCGCAAGATGGGAGACAGAGTGGAAAGCCGGG 503  
593 CCGCGCGAAGAGGCTGTCTGCAGTTCCTGAGTTTGGCGGAATCCA---CCCCGAGCAAAACA 649  
504 CCATTGCCATTGCTGCAGGCTTCAGTGGTCCCTGGCACTAATTCCTCCCATCAATCCC 563  
650 TCGATGAGATCGTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTGTGAAGCGAGTTG 709  
564 TGCATGAGGCACAGAGTTCTTAACACCTATGGTTTCCCTATTATCTTCAAGGCTGCCT 623  
710 CCGGTGGTGGCGGACGGGTATGGTTTGTGTCTTCACTGATGAGCTTCGCAATTAG 769  
624 ATGAGGTGGGCGGTGGCATGAGGTTGTGATAGTACGAGAGCTTGGAGAGATTT 693  
770 CAACAGAGCATCTGTGAAGCTGAAGCGGCTTTGCGGATGCGCGGTATATGTGGAAC 829  
684 ACACCGGCGCTTACCTGAGGCTTGGCAGCTTTGGGAAATGGGCAATTTGTTGGAGA 743  
830 GTGCTGTGATTAACTCCATCAGCATATTGAAGTGCAGATCTTGGCGATCAACATGGAGAAG 889  
744 AATTCAATTGAGAGCAGACACATTTAGGTGCAATCTTAGGGGACCAATATGGGNACA 803  
890 TTGTACACTTTATGAACGTGACTCTCACTGACGCTGCTACACCAAAAAGTTGTGAAA 949  
804 TCTTGCACTTGTATGAGCGGACTGCTCCATCCAGCGCGGACACAGAGTGTGAGAGA 863  
950 TTGCGCCAGCAGCATTTGGATCCAGACTGCGTGTATCGCATTTGTGCGGATGCAGTAA 1009  
864 TTGCCCCGTGTACCCACCTGACACCCCACTTCGGTCAACGCTTACCAGTGACTCTGTCA 923  
1010 AGTTCTGCGGCTCCATTGGTTACAGGGCGGGAAACCGTGGAAATCTTGTGATGAAA 1069  
924 AACTTGCCAAAGCGTTGGCTATGAGATGACGACATGTGGAGTTCTTGTGGACAAAG 983  
1070 AGGCAACACGCTTCTCATGAATGAACCAACGATTCAGGTTGAGCACAACGTTGACTG 1129  
984 ATGCAAGCATACTTCAATCAGGTCATTTCCGCTTCCAGGTTGGAGCACAACGTTCACTG 1043  
1130 AAGAGTCAACGAGTGGACCTGTGTGAAGGCGAGATGCGCTTGGCTGTGGTGCAACCT 1189  
1044 AGGAGATTACAGATGTGGACTGGTTCATGTCTAGATCCATGTCTCCGAGGCGGAGCC 1103  
1190 TGAAGGAATTGGTCTGACCAAGATAGATCAAGACCAACGTTGACGACTGAGTGGC 1249  
1104 TGCCTGACTAGGCTGCGGAGGAAACATCCGAATCAATGGTTGTGCCATTCAAGTGTG 1163  
1250 GCATCAACAGAGATCCAAACAGGCTTCCGCGCAGATACCGGAATATCAACCGCT 1309  
1164 GGGTCAACCACTGAGGACCTTGCACGAGCTTCCAGCCAGACACTGGCGGCAATTGAGTTT 1223  
1310 ACCGCTCACAGGCGAGCTGGCGTTGCTGTGACGGTG---CAGCTCAGCTCGGTGGCG 1366

1224 TCCGAGTGTGTGAGGCGCATGGGCAATCCGCTTGGCAATGCTCAGCATTCAGGAGCTG 1283  
1367 AAATCACCGCACACTTTGACTTCCATGTGTGGTGAATAATGACCTGCGGTGGTTCCCACTTTG 1426  
1284 TCATATCCCCCACTATGACTCCCTGTCTGCTCAAGTCAATGCTCCCAATGGCAAGACCAAC 1343  
1427 AAATGTGTGTGTGTGTCGACAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1486  
1344 CTACAGCTGCCACCAAGATGAGCAGAGCCCTGGCGGAGTTCCGTGTCCGAGTGTAAAGA 1403  
1487 CCACCAATTTGTTTCTGT 1546  
1404 CCACATCTCCCTTCTGTGAGATGTGTCTACACACGAGAGTTCCTAGCGGCGCATTTGG 1463  
1547 CCACCGGATTCATTTGCGGATCACCGGACCTTCTTTCAGGCTTCCACTGTGTGTGTGTGTGT 1606  
1464 ACACCAAGTTCATGATGAGAACCCCGAGCTGTTCAGAGTGTGTGTGTGTGTGTGTGTGTGT 1523  
1607 AGGAGCGCATCTGTGATTTACTTGGCAGATGTCCGCTGAACAAGCCTCATGTGTGTGTGTGT 1666  
1524 CCAGAGT 1583  
1667 CAAAGATTTTTCAGCTCCTATTCGATTAAGC-----TGCCTAAACATCAAGGATCTGCAC 1720  
1584 CGGTCAAGTCTAGTCCAGCCCTGTGGACCCCAATTTCTGTGTGTGTGTGTGTGTGTGTGTGT 1643  
1721 TGCACGCGGTTCCTGTCAGCCGCTGAGAGAGCTTGGCCCGAGCCGTTTCTGTGTGTGTGTGT 1780  
1644 CCCCAGCTGT 1703  
1781 TCCGTGAGCAGGACGACTGT 1840  
1704 TCGGGAATCACCCAGGCGT 1763  
1841 TGTGTGCAACCGAGTCTGCTCATTTGCACTGTGAGCTTCCGCGCAGAGCCGTCGCAAGC 1900  
1764 TACTTGCACATAGATGTGCGCACACACGATCTCAAAAGATTTGACCCCTACGTTTGCACCA 1823  
1901 TGACTCCTGAGCTTTTGTCCGTGTGAGGCTTGGCGCGCGAGCTTACGATGTGTGTGTGTGT 1960  
1824 ACTTCAACAACTCTTTCAGCATAGAGAACTGGGAGGAGCCACATTTGACGTGTGTGTGTGT 1883  
1961 GTTTCCTCTTTGAGGATCCGTGTGGAACAGCTTCGAGAGCTGGGAGGCGGATGCCGATG 2020  
1884 GCTTCTGTATGAGTGCCTTCCGCGCGGCTCCAGAGCTCCGAGGCTCATCCCAACA 1943  
2021 TAAACATTCAGATGT 2080  
1944 TCCCATTCAGATGTCTGTGAGGCGGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2003  
2081 CCGTGTGCGCGGTTTGTAAAGAGTGTGCGAGCTCCGCGGTGAGCATCTTTCGCTGTCT 2140  
2004 AGTGTGTCTTCAAGTCTGT 2063  
2141 TCGAGCGCTTAAACAGCTTCTCCAGATGCGTCCAGCAATCGACGAGTCTGTGGAGACCA 2200  
2064 TTGATCTCTTAACTTACTTGTCCAAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2120  
2201 ACACCGCGTGTAGCCAGGT 2260  
2121 CTGGGCGT 2180  
2261 AGCTTACACCTGT 2320  
2181 CTAAATCTTACTGTGAGTACTACATGGGCTTAGCTGTAAAGAACTGTGTGTGTGTGTGTGT 2240  
2321 ACATCTTGTGCCATTAAGGATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2380  
2241 ACATCTTGTGATTAAGGACATGGCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2300  
2381 TCACCGCACTGG---CCGTGAAATTCGATCTGCCAGTGTGACGTGTGTGTGTGTGTGTGTGT 2437

2301 TCAGCTCCCTCCGGGACCGGTTCCCGAACCCTCCCACTGCCATCCATCCATCCATGACACAT 2360  
2438 CGGGTGGCCAGCTGGCAACTACTTTGCTGACAGCTCAAGCTGGTGCAGATGCTGTTGACG 2497  
2361 CAGGGTCAGGTGGGACGACATGTTGGCTCTGTCACAGCTGGGCTGATGTTGGATG 2420  
2498 GTGCTTCCGACACATGCTGGGACACACTCCAGCCATCCCTGCTGCTGCCATGTTGCTG 2557  
2421 TGCGAGTCACTCTATGCTGGGATGACCTCAAGCCAGCATGGGGGCCCTGGTGGCCT 2480  
2558 CATTCGCGCACACCGCTCCGCTACCGGTTGAGCTCGAGCTGTTCTGACCTCGAGC 2617  
2481 GTACCAAGGAGCTCTCTGGACACAGAGTACCTCCAGGCGTGTGTTGATACATG 2540  
2618 CGTACTGGGAAGCAGTGGCGGACGTGTAACCTGCCATTGA-----GTCTGGAACCCAG 2671  
2541 AGTATTGGGAAGGGCTCGGGGGCTGTATGACGCTTTGATTCACGGCTACCATGAAGT 2600  
2672 GCCCAACCGGTCCGCTTACCGCCACGAAATCCAGGGCGACAGTTGCCAACCTGCGTG 2731  
2601 CTGGCAACTCAGACGTGATGAGATGAGATCCAGGGGCGCAGTACACCAACTACACT 2660  
2732 CACAGCCACCGACATGGGCTTGGGATCGTTTGAACCTATCGAAGACAACTACGACG 2791  
2661 TCCAGGCCACACGATGGGACTTGGCTCCAAAGTTCAAGGAGGTCAAGAGGCTATGTGG 2720  
2792 CGGTTAATGAGATGCTGGGACCGCCCAACCAAGGTCAACCTCCCAAGGTTGTTGGCG 2851  
2721 AGGCTAACAGATGCTGGGGGACCTCATCAAGTGACCACTCCAGATGTTGGGGG 2780  
2852 ACCTCGCATCCACCTCGTTGGTGGGCTGGATCCAGCAGACTTGTGCTGCCGATCCAC 2911  
2781 ATCTGCGCCAGTTTCATGCTGACAGACGGGTTGAGCGGCGGACAGGAGCTCAGGCG 2840  
2912 ABAAGTACGACATCCAGACTCTGTCATCGCTTCTGCGCGGCGAGCTTGTGACCTC 2971  
2841 AAGAGTGTCTTCCCGCGCTCTGTGGTGGAGTTCCTGAGGCTACATTTGCAATTCGCC 2900  
2972 CAGTGGCTGGCAGAGCCACTGCGACCCGCGCACTGGAAGGCC 3016  
2901 ATGGGGGTTTCCCTGAACCTTCCGTTCTTAAGGTGCTAAGGACC 2945

## RESULT 11

US-09-880-107-3029  
; Sequence 3029, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherif, Iwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3029  
; LENGTH: 4017  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 S72370  
US-09-880-107-3029

Query Match 18.7%; Score 677.4; DB 9; Length 4017;  
Best Local Similarity 54.4%; Pred. No. 4.1e-196;  
Matches 1525; Conservative 0; Mismatches 1256; Indels 24; Gaps 7;

QY 233 TCAAAAGATCTTGTGAGCAAAACCGCGGCAAAATCGCGTCCGTGCTTTCCGTGAGCAC 292  
DB 192 TCAAGAAAGTCATGTTGGGCCAACACAGAGGTGAGATTGCCATCCGTGTTTCCGGGCTGCA 251  
QY 293 TCGAAACCGGTGACGACCAAGTAGCTATTATCCCGTGAAGATCGGGATCATTTCCACC 352  
DB 252 CGAGAGTGGGCATCCGACCGTAGCATCTACTCTGAGCAGGACACGGGCCAGATGACCC 311  
QY 353 GCTCTTTTGTCTTGAAGCTGTCCGATTTGGTACCGAAGSCTCACAGTCAAGGGGTACC 412  
DB 312 GGCAGAAAGCAGATGAAGCTTCTCATCGCGCGGCTGCGCCCGCTGACAGGCTACC 371  
QY 413 TCGACATCGATGAATNTCGGTGACGCTTAAAGTTAAAGCAGATGCCATTACCCGG 472  
DB 372 TGCACATCCAGACATCATCAAGTGGCCAGAGAGAACCAACGTAGATGAGTGCACCTTG 431  
QY 473 GATACGGCTTCTCTGTCTGAAATGCCCAGCTTGCCTCGAGTGTGCGGAAACGGCATTA 532  
DB 432 GCTACGGTTCCTCTCTGAGCGAGCGGACTTGGCCCCAGGCTGCCAGGATGACAGGGTCC 491  
QY 533 CTTTTATGGCCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAGTCTCGCGGTTAA 592  
DB 492 GGTATTATGTTCCAGCCCCAGAGGTGTCGCAAGATGGGACAAAGTGGAGGCCCGGG 551  
QY 593 CCGCGCGAAGAGGCTGTGCTGCCAGTTTGGCGG---AATCCACCCCGAGCAAAACA 649  
DB 552 CCATCGCCATTGCTCGGGGTGTTCCCGTGGTCCCTGGCACAGATGCCCATCACGTCC 611  
QY 650 TCGATGAGATCGTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCGATTG 709  
DB 612 TGCATGAGGCCACAGGTTCTCCAAACACCTACCGGCTTCCCCCATCTCTCAAGGCGGCT 671  
QY 710 CCGTGTGTGCGGACGCGGTATGCTTTGTTGTTGCTTCACTGATGAGCTTCGCAATTAG 769  
DB 672 ATGGGGTGGGGCGTGGCATGAGGGTGTGCACAGTACGAGGAGCTGGAGAGATT 731  
QY 770 CAACAGAGCATCTGTGAAGCTGAAGCGCTTTTCGCGATGCGCGGTATATGTGCAAC 829  
DB 732 ACACCGGGCTTACTCAGAGGCTGCGCGGCTTTGGGAATGGGCGCTGTTGTGAGA 791  
QY 830 GTGCTGTGATTAAACCTTCAGCATATTGAAAGTGAGATCTTGGCGGATCACACTGGAGAAG 889  
DB 792 AGTTTCAGAGAGACACCGGCATCGAGGTGAGATCTTGGGGACAGTATGGGACA 851  
QY 890 TTGTACACCTTTTGAAGCTGCTCTACTGAGCGTGTGTCACCAAAAGTTGTGCAAA 949  
DB 852 TCCTGCACCTGTACGAGCGAGACTGCTCCATCCAGCGCGGCGCACCAAGAGGTGTCGAGA 911  
QY 950 TTGCGCCAGCACAGCATTTGGATCCAGAACTGCTGATCGCATTTGTGCGATGCAATA 1009  
DB 912 TTGCCCCCGCGCCACCTGGACCCGAGCTTCGGACTCGGCTCACCGGACTCTGTGA 971  
QY 1010 AGTTTGCCTCCATTTGTTTACCAGGCGCGGAAACCGTGGAAATCTTTGGTCGATGAA 1069  
DB 972 AACTCGCTAAACAGGTGGGCTACAGAAACGAGGACCGTGGAGTTCTTGTGGACAGGC 1031  
QY 1070 AGGCAACACGCTTTCATCGAATGAAACCCAGTATCCAGGTTGAGCACACCGTCACTG 1129  
DB 1032 ACGCAACGACTTCTTATCGAGGTCACTCCGCTGACAGTGGAGCACACCGTCAAG 1091  
QY 1130 AAGAAGTCAACGAGGTGACCTGTGGAAGCGCAGATGCGCTTGGCTGTGTCACACCT 1189  
DB 1092 AGGAGATCACCGACGTAGACTGCTGCTCAGATCCACGTGGCTGAAGGAGGAGGCC 1151  
QY 1190 TGAAGGAATTTGGGTCTGACCCCAAGATAAGATCAAGACCCAGCTGCAGCACTCAGTGCC 1249  
DB 1152 TACCCGACCTGGGCTCGGCGAGAGAAATCCGATCAAGGGTGTGCCATCCAGTGCC 1211  
QY 1250 GCATCACCGGAAGATCCAAACAGGCTTCCGCCAGATACCGGAACATATCACCGCGT 1309  
DB 1212 GGGTCACACCGAGGACCCCGCGCCACGTTCCAGCGGACACCGGCGCGCATTTGAGTGT 1271

QY 1310 ACCGCTACAGGGCGAGCTGGGCTTCGTCT---TGAGGTGACAGCTCACTCGGTGGGG 1366  
DB 1272 TCCGAGCGGAGAGGGCATGGGCTTCGGCTGGATAATGCTTCGGCTTCCAGAGGACCG 1331  
QY 1367 AAATACCGGACACTTTGACTCCATGCTGTGGTAAATGAAGCTGCGGTGGTTCGAGCTTTG 1426  
DB 1332 TCACTCGCCCACTACGACTCCCTGCTGGTCAAGTCAATGCCCAGGCAAGACCAACC 1391  
QY 1427 AAATGCTGTGCTGCTGCAAGCGCGGTGGTGGTGAATTCACGCTGCTGTGGTGTGCA 1486  
DB 1392 CCAGCGCGCCACCAAGATGAGAGGGCCCTTGGCGAGTTCCGCTGCTCGAGGTGTGAAGA 1451  
QY 1487 CCAACATTTGTTCTTGGCTGGTGTGCTGGGGAAGAGGACTTCACTTCCAAAGCGATCG 1546  
DB 1452 CCAACATCGCTTCTGAGAAATGTGTCAACACAGAGCTTCTGGGAGGCACTGTGG 1511  
QY 1547 CACCGGATTCATGTCGATCAACCGGACCTCTCTTCAAGGTTCACCTGCTGTGATGAGC 1606  
DB 1512 ACACCCAGTTTCATGAGAGAACCCAGACGTGTTCCAGCTGCGGCTTGCACAGAACCGGG 1571  
QY 1607 AGGAGCGCATCTGATTTACTTGGCAGATGTCACGCTGMA---CAAGCTCATGTTGTG 1662  
DB 1572 CCCAAAGCTGTGACTTACCTGGCCATGTGATGTTAAAGTTCACACACCCGATTC 1631  
QY 1663 CGTCCAAAGATGTGACGTCTTATCGATAAGCTG---CCTAACATCAAGATCTGCCAC 1720  
DB 1632 CGTCAAGGCCAGCCACGACCCACGAGACCCCGTGTCTGCTGCACTGCCATAGGCCCGC 1691  
QY 1721 TGCACAGGGTTCCTGCTGACCGCTGAAGAGCTGCGCCAGCGCTTGTCTGCTGATC 1780  
DB 1692 CCCCGGCTGTGTCAGAGACATCTGTGCGAGAGGGGCTGAGGGCTTGTCTGAGCTG 1751  
QY 1781 TCCGTGAGCAGGAGCACTGAGCTTACTGATACACCTTCGCGATGACACACAGCTTT 1840  
DB 1752 TGGGAAACACCCGCGGCTGTGCTGATGGACACGACCTTCAGGAGCGCCACCACTCAC 1811  
QY 1841 TGCTTGCAACCGAGTCGCTCATTCGACATGAGCTGCGGAGAGCGCGTGCACAGC 1900  
DB 1812 TGCTGGCCACTGCTGTGCGCACCCACCGATCTCAAAAAGATGCCCCCTATGTTGCCACA 1871  
QY 1901 TGACTCTCTGAGCTTTGCTCGGTGAGGCTCGGGCGCGGACCTACGATGTGGCGATGC 1960  
DB 1872 ACTTCAGCAAGCTCTTCAGCATGAGAACTGCGGAGGAGCCACGTTTGAAGTCCGCTATC 1931  
QY 1961 GTTTCCTTTGAGATCCGTGGGACAGCTCGAGAGCTCGCGAGGCGATGCCGATG 2020  
DB 1932 GCTTCCTGTATGAGTCCCTTGGCGGCGCTGCAAGAGCTCCGGAGCTCATGCCCAACA 1991  
QY 2021 TAAACATTCAGATGCTGCTTCGCGCGCAACACCGTGGGATACACCCCGTACCCAGACT 2080  
DB 1992 TCCGTTTCCAGATGCTGTGCGGGGGCAATGCTGTGGGCTACACCACTACCCAGACA 2051  
QY 2081 CGCTCTCGCGCGCTTTGTTAAGGAAGTGCAGCTCCGGGTGGACATCTTCCGCACT 2140  
DB 2052 AGCTGGCTTCAAGTCTCTGAAAGTGGGCAAGAGATGGCATGATGCTTCTCGGTGT 2111  
QY 2141 TGGAGCGCTTAAACAGCTTCCAGATGCGTCCAGCAATCGAGCACTCTCGAGACCA 2200  
DB 2112 TTGATCTCCTCACTACTTGGCCACATGCTGCTGGGATGAGGCGG---CAGAGTG 2168  
QY 2201 ACACCGCGGTACCGAGGTGGCTATGGCTTATTTCTGATGATCTCTGATCCAAATGAAA 2260  
DB 2169 CGGAGGGGTGGTGGAGGCTGCATCTCATACACGCGGCGAGTGGCCCGACCCAGCGCGCA 2228  
QY 2261 AGCTCTACACCTGATTTACTTACCTAAAGATGGGAGGAGATCGTCAAGTCTCGGCTC 2320  
DB 2229 CCAAGTACTCTACGAGTACTACATGGGCTTGGCTTGAAGCGCTGTGCGAGCTGGACCC 2288  
QY 2321 ACATCTTGGCCATTAAAGATATGGCTGTCTGCTTCCGCCAGCTGCGGTAAACCAAGCTGG 2380  
DB 2289 ACATCTGTGATCAAGGACATGGCGGCTGCTGAAGCCACCGCTGCTGACCATGCTGG 2348  
QY 2381 TCACCGCACTGG---CCGTGAATTCGATCTGCCAGTGCACGTCACACCCACGACACTG 2437

DB 2349 TCAGCTCCTCTCGGGACCGCTTCCCGACTCCCACTGCAATCCACACCCACGACCGT 2408  
QY 2438 CGGTGGCCAGCTGGCAACCTACTTTGCTGAGCTCAAGCTTGGTGGATGCTGTGACG 2497  
DB 2409 CAGGCGCAGGGGTGGCAGCCATGCTGGCTGTGCCAGGCTGGAGCTGATGTGGTGGATG 2468  
QY 2498 GTGCTTCGCGACCACTGCTGSCACCACTCCAGCCATCCCTGCTGCTGCCATTTGTGCTG 2557  
DB 2469 TGGAGCTGATTCATGCTGGATGATCTTCAAGCCAGCATGGGGCCCTGTGGGCT 2528  
QY 2558 CATTCGCGCACACCCGCTCGCATACCGGTTTGAAGCTCGAGGCTGTTTCTGACTCGAGC 2617  
DB 2529 GTACAGAGGAGCTCCCTCGACACAGAGGTGCCCATGGAGCGGCTGTTTACTACAGTG 2588  
QY 2618 CGTACTGGAGAGCTGCGCGGACTGTACTCTGCAATTTGAGTCTGGAACCCACGCCCCAA 2677  
DB 2589 AGTACTGGAGGGGCTCGGGGACTGTACGGGCTTTCGACTGCAAGCCACCATGAAGT 2648  
QY 2678 CCGG-----TGCGCTTACCGCCACGAAATCCAGGCGGACAGTTGTTCACCTGCGTG 2731  
DB 2649 CTGGCAACTCGGACGCTGTATGAAATGAGATCCAGGGGGCCAGTACACCAACCTGCACT 2708  
QY 2732 CACAGGCCACCGCACTGGGCTTTCGGATCGTTTGAATCATCGAAGCACTACGACG 2791  
DB 2709 TCCAGGCCCCACAGCATGGGGCTTGGCTCCAAAGTTCAAGAGGTCAAGAGGCTATGTGG 2768  
QY 2792 CCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCAACCCATCTCTCCAAAGTTTGTGGCG 2851  
DB 2769 AGGCCAACCAAGATGCTGGGCGATCTCATCAAGGTGACGCCCTCTCCAAAGATCGTGGGG 2828  
QY 2852 ACTCGCACTCCACTCGTTGCTGGGTGGATTCAGAGCACTTTGCTGCCGATCCAC 2911  
DB 2829 ACCTGGCCCCAGTTTATGTTGCAAGATGAGTTGAGCCGGGCGAGGCCGAAGCTCAGGGCG 2888  
QY 2912 AAAAGTACGACATCCAGACTCTGTCATCGCTTCTCGCGGGCGAGCTTGGTAAACCTC 2971  
DB 2889 AAGAGCTGCTCTTCCCGCTCCGTGGTGGATCTCTGAGGCTACATCGGTGCCCC 2948  
QY 2972 CAGGTGGCTGGCAGAGCACTGCGCACCCGCGCACTGGAAGGCC 3016  
DB 2949 ATGGGGGTTTCCCGAACCCCTTCTGCTTAAGGTACTGAAGGACC 2993

## RESULT 12

US-10-369-493-31453  
; Sequence 31453, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; PRIOR FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 31453  
; LENGTH: 3432  
; TYPE: DNA  
; ORGANISM: Rhodobacter sphaeroides  
US-10-369-493-31453

Query Match 18.2%; Score 659; DB 15; Length 3432;  
Best Local Similarity 51.8%; Pred. No. 1.6e-190;  
Matches 1782; Conservative 0; Mismatches 1580; Indels 75; Gaps 10;

QY 238 AAGATCTTGGTAGCAACCGCGCGGAATCGCGGTCCGTGCTTTCGTCGAGCACTCGAA 297

1 AAGATCCTCATGCCAACCGTGGTGAATCGCATCCGGGTGATCGGGCGGCAACAGAG 60  
298 ACCGGTGCAGCAGCGTACGATATTTACCCCGTGAAGATCGGGGATCATTTCCACCGCTCT 357  
61 ATGGGCAAGAGACGGTCCGCTCTATGCCGAGGAGACAGACTCTCGCTCCACCGCTTC 120  
358 TTGTCTTCTGAAGCTGTCGCGATTTGGTACCGAAGGCTCACAGTCAAGGCGTACCTGGAC 417  
121 AAGGCGGACAGGCGCTTACCGGATCGGCGAGGCGCTCTCGCCCGTGGGCGCTATCTCTCG 180  
418 ATCGATGAATATCGGTCCAGTCTAAAGATTAAGCAGATGCCATTTACCCGGGATAC 477  
181 ATCCCGAGATCATCCGCTGGGCGAGATGTCCGGCGCGATGCCATCCACCGGCTAC 240  
478 GGCTTCTCTGAATAAGCCAGCTTTGCCCGGAGTGTGGGGAAGAAACGGCATATCTTTT 537  
241 GGGCTTCTCTCGAGAAACCCGATTTCTGTCGAGGCGCTCGACGACGAGCGGCGATCGCCTTC 300  
538 ATTGGCCCAACCCAGAGGTCTTGATCTCACCGGTGATTAAGTCTCGCGCGGTAAACGCC 597  
301 ATCGGCGGAGAGGCGGCGGAGATGCGCGGCTCGGCGACAGGCTTCGCGCGCGCGTG 360  
598 GCGAAGAGGCTGTGTGCGATTTTGGCGGAATCC---ACCCGAGCAAAACATCGAT 654  
361 GCCATGGCGCGGCGGTGCGGTGATCCCGGCGACCGAGGTCTCGGCGACGATATGGAG 420  
655 GAGATCGTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAGGCGAGTTCGCGT 714  
421 GAGATCAAGCGGAGGCGGCGGAGATCGGCTATTCGCTGATGCTCAAGGCTCTGTGGGCG 480  
715 GGTGGCGGACGCGGTATGGCTTTTGTGCTTCACTGATGAGCTTCGCAAAATAGCAACA 774  
481 GCGGTGGCGGCGGATCGGCCATCACTTCGAGGCGAGCTGGCGGACAAAGTGGCG 540  
775 GAAGCATCTCGTGAAGCTGAAGCGGCTTGGCGATGGCGGATATGTCGAAGCTGCT 834  
541 GAGGCGGACGCGAGGCGCGGCGCTTTCGGAACGCGGCGGCTATCTCGAAGATG 600  
835 GTGATTAACCTCAGCATATTGAAGTGCAGATCTTGGCGATCACACTGGGAAGTTGTA 894  
601 ATCCAGCGCGCGGCGCACGTGAGGTGCAGATCTCTCGGCGACAAATACGGCGCATCTAC 660  
895 CACCTTTATGAAGTGAAGTCACTGAGCGCTGCTCACCAAAAGTTGTGCAAAATGGG 954  
661 CATCTCTACAGGCGGAGTGCACCGTGCAGCGCGCAACGAGAGGTGTGAGCGCGG 720  
955 CCAGCAGCATTTGGATCCAGAACTCGGTGATTCGCAATTTGTCGGATGCAATAAGTTTC 1014  
721 CCGCGCTTATCTACCGGAGGACGCGACAGAGATCTCGAGCTCGGCGCGCGATC 780  
1015 TGCCGCTCCATTGGTTACCGGGCGCGGAAACGTTGGAATTTCTGGTCGAT---GAAAG 1071  
781 TGGCGCCCATGTGAATTACGAATGCGGCGGACGGTCTGAGTTCTCTGATGATATGAGCTCG 840  
1072 GGCAACACAGCTCTTATCGAAATGAACCAACGATATCCAGTTGAGCACACCGTGAAGTAA 1131  
841 GAGAAATCTATTCTATCGAGGTCAATCCGCGGTTTCAGTTCGAACATACCTGACCGAG 900  
1132 GAATGACCGAGGTGAGCTGTGTAAGGCGGAGATGCGCTTGGCTGTGCTGGTCAACCTTG 1191  
901 GAAGTGAAGGCGCATGACATCGTGCAATCCAGATCCGATCCGATCCGGAAGGCGGCGGTTG 960  
1192 AAGGAATTGG---GTCTGACCCAGATTAAGTCAAGACCCAGGTCGAGCATCTGCGAG 1245  
961 GCGGAGGCAACCGGATGCGCGAGCGAGGAGCATCAAGCTCTCGGCGCATGGCTGCGAG 1020  
1246 TGGCGGATACCAACGGAAGATCAAAACAGGCTTCCGCGCGAGATACCGGAATATACCC 1305  
1021 TGGCGGCTACGACCGGAGATCCGCGAAGCAATTTCAATCCGAGTACGCGCGGCTCACCC 1080  
1306 GCGTACCGCTCACAGGCGGAGTGGCTTCTGCTTGAAGGCTG---CAGCTCAGCTCGGT 1362

1081 GCCTATCGTCTGGCCACGCGGCGATGGGCGATCCGGCTGGAGCGGCGGCGATATGCCGGG 1140  
1363 GCGAAATACACCGCACACTTTGACTCTCATCTGCTGTAATAATGACCTGCGGTGTTCCGAC 1422  
1141 GCGGTATACACCGCTATTACGATTCACCTCTCGTGAAGGTGACGCGCTGGGCGCCACG 1200  
1423 TTTGAAATCTGCTGCTGTCGACAGCGCGGTTGGCTGAGTTACAGTCTCTGCTGTT 1482  
1201 CCGGAAAGGCGCATCGCGGCGATGACCGGGCGCTGCGGAGTTCGCGATTCGCGGCGTG 1260  
1483 GCAACCAACATTTGCTTTCTTGGCTGCTGCGGGAAGAGACTTTCATCTTCCAAGCGC 1542  
1261 GCCACCAACATGCGCTTCGTCGAGAACCTGCTGAAGCACCGGAGTTTCTCTGACTATTCC 1320  
1543 ATCGCACCGGATTCATTGGCGATCACCGGACCTCTTCAGGCTCCACTGCTGATGAT 1602  
1321 TACACGACGAATTCATCGACGACGCGCGGACCTTCACTTCAAGCGCGTCTGAC 1380  
1603 GAGCAGGACGCACTCTGATTTACTTTGGCAGATGTCAACGTTGAAACGCTC-----AT 1656  
1381 CCGGCGAGGAGATCTGACCTACATCGCGGACATCACGTTGAACGGGCAACCGCGAGCG 1440  
1657 GGTGTGGTCCAAAGGATGTTGAGCTCTATCGATGAGTCTTAACATCAAGGATCTG 1716  
1441 GCGGCGCGGTCCGCGCTTCGCGCGAGCTGAAGGATCCGAAGGCAACCGAGGCCAAGGCG 1500  
1717 CCACTGCCACGCGGTTCCCGTGACCGCTGAAGCAGCTTTGGCCACGCGCTTTGCTCGT 1776  
1501 GCAACCGGCGCGGACGCGGACCTCTCTGAGGAGAGGCGCGGCGGTGGCGAC 1560  
1777 GATCTCGGTGAGCAGGACGACATGGCAGTTACTGATACCACTTCGCGGATGACACAG 1836  
1561 TGGATGCGCGCGCAGACCGCGCTGCTGATGACGACGACGATGCGGACGCGCACCG 1620  
1837 TCTTTGCTTCGACCGCGAGTCCGCTCATTTGCGACTGAAGCTTCGCGCAGAGGCGCTCGCA 1896  
1621 AGCTGCTGCGCACCGCGATGCGCTGATGCGATGATCAAGGTCAAGGTCAAGGTGATGCG 1680  
1897 AAGTGAATCTGAGCTTTTGTCCGTGGAGGCTGCGGCGGCGGACCTTACGATGTTGGG 1956  
1681 GCGAAACCTCGCGCGCTCTTCTCGTGGAAATGCTGGGCGGCGGCGCACCTTCGATGTGGCC 1740  
1957 ATGCTTTCTCTTTGAGGATCGTGGCAGCTCGACGAGCTGCGGCGGCGGAGGTGCGG 2016  
1741 TATCGCTTCTCGAGGAATGCTCTCGGAGCGGCTGCGGACATCGCGCGCGGCTGCCC 1800  
2017 AATGTAAACATTGAGATGCTTTCGCGCGCGCAACACCGTGGGATACACCGCGTACCCA 2076  
1801 AAGTCAATGACGAGATGCTGCTGCGCGCTCGAAACGCGGCTCGGCTATACGAATACCCC 1860  
2077 GACTCCGCTTCGCGCGGTTTGTAAAGGAGCTGCCAGCTCCGCGGTGGACATCTTCCG 2136  
1861 GACAATGTGTGAGGAATTCGTCGCTCAGGCGGCGGAAACCGGCGTGCAGCTGTTCCG 1920  
2137 ATCTTCGACCGGCTTAACGAGCTCTCCAGATCGTCCAGCAATCGACGAGTCTTGAG 2196  
1921 GTGTTGACTCGCTCACTCGGTGAGAACATCGCGGTGGCAATGATGCGGTGATCGAG 1980  
2197 ACBAACCGCGGTAGCGGAGGTGGCTATGGCTATTCTGGTATCTCTGATCCAAAT 2256  
1981 GCCAAACA---AGGCTCGGAGGCGACGATCTGCTACGCGGCGACCTGTCTGATCCGAC 2037  
2257 GAAAGCTCTACACCCCTGGATTTACTACCTAAAGATGGCAGAGAGATCGTCAAGTCTGGC 2316  
2038 CGCTCCAAATACGATCTGAATCTACTATGTCGGCATGCGGCGCGCTGCGCGAGCGGGC 2097  
2317 GCTCAGATCTTGGCCATTAAAGGATATGGCTGCTGCTTCCGCCAGTGGGTAAACCAAG 2376  
2098 GCGCATGTGTGGGTGTGAAGGACATGCGCGGCTTCTGAAGCCTTCCCGCGCGCGGCTG 2157  
2377 CTGTGTCACCGCACTGCGCGCTGAAATTCGATCTGCGCATGACGCTGCACACCCACGACCT 2436  
2158 CTGTGTAAGGCGCTGAAGGAGGAGGTGGGCTCCCGATCCACTTCCACACCCGACGACG 2217

QY 2437 GCGGTGGCCAGCTGGCAACTACTTTGCTGCGAGCTCAAGTGTGTGTCAGATGCTGTGAC 2496  
Db 2218 AGCGGAATTGGGGCGGACCGTGTCTCGCGCTCGCGATGATCGGGGTGATCGGTGAT 2277  
QY 2497 GGTGCTTCCGACACACTGTCTGGACACACCTCCAGGCATCCCTGTCTGCCATTGTGCT 2556  
Db 2278 GCGGCGATGGAGCGCTTCTCGGGCGGACACTCGAGCCCTGCTCGGCTCGATGCTGAG 2337  
QY 2557 GCATTGCGGCGACACCGTTCGCGATACCGGTTTGAAGCTTCGAGGCTGTTTCTGACCTCGAG 2616  
Db 2338 GCGCTGAAGCACACCGACCGCGACACGCGCTCGACATCGGGCGATCCGCGAGATCTCG 2397  
QY 2617 CCGTACTGGGAAGCAGTGGCGGAGTGTACTGTGCTGATCTGAGTGTGAACCCCGAGCCCA 2676  
Db 2398 GACTACTGGGCGCATGTGCGCCAGCAGTATTCGGCTTCGATCGGCCCTGCGCCCTCGCCC 2457  
QY 2677 ACCGCTCGCTTACCGGCGACGAATCCAGCGGCGAGTGTCCAACTCGCGTGACAG 2736  
Db 2458 GCCTCCGAGTCTATCTGACAGATGCGCGGCGCCAGTTTCAACACTCAAGGCGCAG 2517  
QY 2737 GCGACCGCACTGGGCTTGGCGATGCTTTCGAACTCATCGAAGACAACTAGCGAGCGGTT 2796  
Db 2518 GCGGCTCCATGGGCTCGAGGAAGCTGTTCGAGGTGGCGAGGCTATGCCGAGCGG 2577  
QY 2797 AATGAGATGCTGGGACGCCCAACCAAGTCAACCCATCTCCAAAGTGTGGGACCTC 2856  
Db 2578 AACCGAATGTCGGGCAATCGTGAAGTCAAGCTCTGCTCGAAGTGTGGGCGAATG 2637  
QY 2857 GCATCCACCTCGTTGGTGGCGGTGTGATCAGCAGACTTTGTCGCGATCCCAAAAG 2916  
Db 2638 GCGCTGATGATGTTGGCGCAGGGCTGACGCGGGAAGAGTTCGAGGATCCGAGGTGAG 2697  
QY 2917 TACGATATCCAGACTGTGATCGGTTCTGCGCGGAGCTTGTAACTCCCTCCAGGT 2976  
Db 2698 GTGAGCTTCCGGAATCGGTGTGACATGCTCAAGGCAATCTCGGCGAGCCCAAGC 2757  
QY 2977 GGTGGCGAGAGCCACTCGCACCGCGCACTGGAAGCGCCCTC----- 3020  
Db 2758 GGTGGCGAGCGATCTCTGAAGAAGTGTGAAGGCGAGGACCTCTCGACCGAGCGA 2817  
QY 3021 -----CGAAGCAAGCACTCTGACGGAAGTTCCTGAG 3054  
Db 2818 CCGGGCGGCAATCTCGCGCGCTGACATCGAGCGCGCGCGAGAGCTCTCTGCGAG 2877  
QY 3055 GAAAGAGCGGCGCACTCGAGCTGATGATTCGAAGGAACGTGCAATAGCTCAACCG- 3113  
Db 2878 ATCAAGCAGGCGACGACGATCCGCTCGACACGCGCGTGCATGCGAGATCTGAACGGC 2937  
QY 3114 --CCTGCTGTTCCGAAGCAACCGAGAGTTCCTGAGCACCGTTCGCGCTTCGGCAAC 3171  
Db 2938 TACCTCATGTATCCCAAGTCTTCAACGACTACCGCGCCGCCACAGGATCTACGGGCGG 2997  
QY 3172 ACCTCTGCGCTGATGATCGTGAATTTCTTACGGCTGTGTCGAAGCGCGGAGACTTTG 3231  
Db 2998 GTGCGGACGCTGCGACCGGACCTTCTTACGGGATGGAGCGGCGGAGAGATCTCG 3057  
QY 3232 ATCCGCTTCCGAGATGTGGGCAACCCACTGCTGTTGTTGCTGATGATGATCTGAGGCA 3291  
Db 3058 GCGGAAAT---CGACCCCGGCAAGACGCTCGAGATCCGGCTGTCCGCGGTGGGCGAGACG 3114  
QY 3292 GACGATAAGGATGCGCAATGTTGTTGGCAACGTCACGCGGCGAGATCCGCGCAATGCT 3351  
Db 3115 TCGGACGAGCGGATGCGCAAGTGTCTTTCGAGTGAACGCGCGAGCGCGGCTGATCCG 3174  
QY 3352 GTGCGTGAACGCTCGTGTGAGTCTGTCAACCGCAACCGGAGAAAGGCGAGATCTCTCAAC 3411  
Db 3175 GTGGCAACCGGCGGTGAAGGCGCAAGACCGGCAACCGGCGCAAGGCGGAGGCGCAAT 3234  
QY 3412 AAGGCGCATGTTGTCGACCATTTGCTGTTGTCACCC---GTGACTGTTGCTGAAGGT 3468  
Db 3235 CCGGCGCATGTCGGCGCCCGGATGCGGCTTCGCTGCGGTCTCGGCGGCG 3294

QY 3469 GATGAGGTCAAGGCTGAGATGTCAGTTCGAATCATTCGAGGCTATGAGATGGAAGCAACA 3528  
Db 3295 CAGAAGGTGAAGCGCGGCGATCTCTCTGTCACCATCGAGGCGATGAAGATGAGACCGGG 3354  
QY 3529 ATCACTGCTTCTGTTGAGCGCAAAATCGATCGCGTGTGTTCTGCTGCAACGAAGTGT 3588  
Db 3355 CTCATGCGACCGTTCGCGCACCGTGAAGCTGTCCATGTGCGGCGCGCGCGAGATC 3414  
QY 3589 GAAGTGGCGACTTCAT 3605  
Db 3415 GAGGGAAGGACCTGCT 3431  
RESULT 13  
US-10-369-493-35714  
; Sequence 35714, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 35714  
; LENGTH: 3459  
; TYPE: DNA  
; ORGANISM: Mesorhizobium loti  
US-10-369-493-35714  
Query Match 17.8%; Score 643.6; DB 15; Length 3459;  
Best Local Similarity 51.6%; Pred. No. 8.7e-186;  
Matches 1788; Conservative 0; Mismatches 1599; Indels 75; Gaps 11;  
QY 229 GCATTCAAAAGATCTTGTTAGCAAAACCGCGCGAAATCGCGTCCGTGCTTTCGTGCA 288  
Db 4 GCATCAAGATCTCTGTCGCGCAACCGTCCGATCAGAAATCGCCATCGCGTCTTTGCGGG 63  
QY 289 GCATCAAAACCGGTGACGCCACGCTAGCTATTTACCCCGGTGAAGATCGGGATCATTC 348  
Db 64 GCCAATGAATCGGCGCTCAAAACCGTGGCGATCTGGCGCGAGGAAGACAAATATTGCTG 123  
QY 349 CACCGCTCTTTGCTTCTGAAGCTGTCGATTTGTTAC-----GAA 390  
Db 124 CACCGCTTCAAGCGCGACGAGAGTACACAGTTCGGGCGGCTCGCATCTCAGAGGAT 183  
QY 391 GGCTCAACAGTCAAGCGGTACCTGGACATCGATGAATTTATCGGTGCAAGTAAAGATT 450  
Db 184 ATGGGCGCGATCGAGAGTATCTGTGATCGAGGAGTGTATCGCGTTCGCGGCTCTCG 243  
QY 451 AAGCAGATGCCATTTACCGGCGATACCGCTTCTGCTGAAATTCGCCAGTTCGCCGCG 510  
Db 244 GCGCGCGATGCCATTCACCCCGGCTACCGGCTCTGTGCGAAAGCGCGGAAATTCGCCG 303  
QY 511 GAGTGTGCGGAAACCGGCAATTTATTTGCGCCAAACCCAGAGGTTCTTGTATCTCAC 570  
Db 304 GCCTGCGCGCAAGCGGCGATCACCTTCATCGGCGCGAGCGGACACGATGCGCGCTC 363  
QY 571 GGTGATAGTCTCGCGCGTAAACCGCGCGGAGAGAGGTGTGCTGCCAGTTTGGCGGA 630  
Db 364 GGCAACAGGTGCGCGCGCGCAATCTCGCATCGAGGTTCGGCTGCTCATTCGCCGCG 423  
QY 631 TCCACCC---GAGCAAAACATCGATGAGATCGTTAAAGCGCTCAAGGCCAGACTTAC 687  
Db 424 ACCGATCGGTTGCCGACCATGATGAGCGGCTCAAGAACTGCGCGAGAGATCGGCTAT 483







1174 GCTGCTGGTCAACCTT-----GAAGGATTTGGTCTGACC---CAAGATAAGATCAAG 1224  
Db  
967 CTCGAAGGAGCGGCCATCGGCACGCGGAATCCGCGTSCCAAGCAGGAAGATATCCG 1026  
Qy  
1225 ACCACGGTGCAGCATGTCAGTGCCTGATCACCACGGAAGATCCAAACAAACGGCTTCGCG 1284  
Db  
1027 CTCACGGCCATCGCTGTCAGTGCCTGATCACCACGGAAGATCCGGAACAAATTTCAAT 1086  
Qy  
1285 CAGATACCGGAACATACCGGTACCGCTCACAGGGGGAGCTGGGGTTCGTTCGAC 1344  
Db  
1087 CCGGACTATGCGCGTATACCGCTATCGCTCGGTTCCGGTTTCGGCATCCGCTCGAT 1146  
Qy  
1345 GGTG---CAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCATGCTGGTGAAA 1401  
Db  
1147 GGGGGCACCTCTATACCGCGCGCTCATCACCGCTTATTAGATCCGCTGCTCAAG 1206  
Qy  
1402 ATGACTGCGGTGGTTCGCACTTTGAACTGTGTGCTGTCGTCGACAGCGCGGTGCT 1461  
Db  
1207 GTCACGGCTGGCGCGCGAGCGGATGAAGCGATCAGCCGATGATCGCGCTTCGCG 1266  
Qy  
1462 GAGTTCACCGTGTCTGCTGTGCAACCAACATTTGGTTTCTTCGCTGCGTGTGTCGGGAA 1521  
Db  
1267 GAATTTCTGATCCGCGCGTTCGACCAACCTGACCTTCTCGNAGCCATCATCGTCA 1326  
Qy  
1522 GAGGACTTCACTTCCAGGCGATTCGCCACCGGATTCATTCGCGATCAACCGCACCTCTT 1581  
Db  
1327 GACAGTTTCGCAACAAACCTATACGACGCTTTCATCGATTTCGACGCGGAGTTGTT 1386  
Qy  
1582 CAGGCTCCACCTCTCATGATGAGCAGGAGCATCTCTGATTAATTGGCAGATGTCACC 1641  
Db  
1387 GCGCAGGTCAAGGTCAGACCGCGCCACCAAGCTTCTGACCTATCTGGCGATGTACC 1446  
Qy  
1642 GTGAACAAGCTCATGGT-----GTGCTCAAGGATTTGTCAGCTCTCTATCGATTAAG 1695  
Db  
1447 GTCAATGGTCAACCAAGAAACCAAGGGTCTGGGAAGCTTCGCAACAGCGCGCAAGGCC 1506  
Qy  
1696 CTCGCTAAACATCAAGATCTGCCACTGCAACGCGTTCGCGTACCGCTGAAGCAGTT 1755  
Db  
1507 ATGCTGCCCTATATCATGTCGCGAGCTCTGAGGACCAAGCAATTCGTGCAAGTT 1566  
Qy  
1756 GGCCACGCGCTTCTCGTATCTCCGTGACGAGCAGCACTGGCAGTTACTGATACC 1815  
Db  
1567 GGCCTTAAGGGCTTTCGGAGTGGATGCGCAACGAAAGCGGCTGCTGTCACGGATACG 1626  
Qy  
1816 ACCTTCGGGATGACACACAGCTTTGCTTGGACCCGAGTCGCTCATTCGACATGAAG 1875  
Db  
1627 ACCATGCGCGAGCAGACAGTCTGCTGGCCACCCGCTCCGTACCATCATGATTCGCC 1686  
Qy  
1876 CCTGCGGAGAGCCGTCGCAAGCTGACTCTCTGAGCTTTTGTCCGTGGAGGCTGGGCG 1935  
Db  
1687 CGGTTTCCAGCGTTTATTCGAAGGCGCTTCGCACTTCTGCTGGAATGCTGGGCG 1746  
Qy  
1936 GCGCGACCTACGATGGGAGTGGTTTCTTCTTTGAGATCCGTGGACAGGCTCGAC 1995  
Db  
1747 GGTGGACCTTTCAGCTCTCAATGCGCTTCTTCGACCAAGACCCGTGGAGGCGCTGTCA 1806  
Qy  
1996 GAGCTCGGAGGCGATCGCAATGTAACATTCAGATGCTGCTTCGCGCGCGCAACACC 2055  
Db  
1807 CTCATTCGGAAGGTGCGCGAACCTGCTGTTGTCAGATGCTTCTGCTGTCGCAAGCGC 1866  
Qy  
2056 GTGGATACACCCGTTACCCAGACTCCGTCTCGCGCGGCTTTGTTAAGAACTGCGCAGC 2115  
Db  
1867 GTCGGTTACAAAGAACTATCCCGATAATGTCGTGAAGTATTTCTGTCGTCAGCGCAAGG 1926  
Qy  
2116 TCCGGGTGGACATCTTCGCACTTCGACGCGCTTAAAGAGTCTTCCAGATGCGTCCA 2175  
Db  
1927 GCGGGCTCGATCTTTTCGCTTTCGACTGCTGAACTGCGGTGGAGAATATCGGGTG 1986  
Qy  
2176 GCAATGACGCACTCTGAGACCAACACGCGGTAGCGAGTGGCTATGCTTATCT 2235  
Db  
1987 TCGATGATCGGATGTCGCGAGGAGNACA---AGCTCTGCGAGGCAACCATCTGCTACACC 2043

2236 GGTGATCTCTCTGATCAAAATGAAGCTCTACACCTGGATTACTTACCTAAGATGCCA 2295  
Db  
2044 GCGATCTCTTGAATTCGGCAGCTCGAAATATGATCTCAATATTTATACCAACCTCGCG 2103  
Qy  
2296 GAGGAGATCGTCAAGTCTGCGCTCATCTTTGGCCATTAAGGATATGCTGCTGCTT 2355  
Db  
2104 GTCAGCTGGAAGAGCGCGGCCATATCATCGCGCTCAAGGATATGCGGCGCTGCTG 2163  
Qy  
2356 CGCCAGCTGCGGTAAACCAAGTGTGTCACCGCACTGCGCGTGAATTCGATCTGCCAGTG 2415  
Db  
2164 AAACCGCGCGCCCAAGGTTCTTTCAAGGCTTGGCGAGCGACCGTCTTTCGCGCGTGGAT 2223  
Qy  
2416 CAGTGCACACCCACGACACTCGGTCGCGCACTGCGCAACCTACTTTGCTGCACTCAA 2475  
Db  
2224 CACTTCCACACCATGATAGTTCGCGCAATTCGCGCGACCGCTTCTTTCGCGCGTGGAT 2283  
Qy  
2476 GCTGCTGAGATGCTGTTGACGCTTTCGCGCACTGCTGCGCACTGCTGCGCACTCCAGGCA 2535  
Db  
2284 GCGGGCTGACCGCTTCGATGCGCGATGATGCTTCTTCGCGCAATGATGCTGCGAGCCC 2343  
Qy  
2536 TCCCTGTCTGCCATTTGCTGCTTTCGCGCACTCCGCTCGCGATACCGGTTTGAAGCTC 2595  
Db  
2344 TGCTCGCTCGATCGTGGAGCGCTCTCGGATCGAAACGCGATACGGTCTCTGATACC 2403  
Qy  
2596 GAGCTGTTTCTGACCTCGAGCGCTTACGCGAGCAGTGGCGGCTGCTGATGCTGCACTATC 2655  
Db  
2404 GAATGATTCGCGCATTTCTTCTTATTTGGAAGCGTGGCAATCAATATGCGCGCTTC 2463  
Qy  
2656 GAGTCTGGAACCCAGCGCCCAACCGCTCGCTCTACCGCACGAAATCCAGCGGACAG 2715  
Db  
2464 GAGATGACCTCAAGGGCTGCTCGGAAGTCTATCTCATGAATGCGGGTGGCCAG 2523  
Qy  
2716 TTGTCACCTCGCTGTCACAGGCACTGCGCTTGGGCTTGGGATGCTTTCGATCTCATC 2775  
Db  
2524 TTCAACCACTCAAGGAACAGGCACTTTCGCTGGTCTCGAAAGCGCTGGCAAGGCTG 2583  
Qy  
2776 GAAGAACAATTCACGAGCGCTTAATGAGATGCTGGAGCGCCCAACCAAGGTCAACCCATCC 2835  
Db  
2584 GCGAGGCTTATCGGATGCCAAGGATTTTCGGGATATCGTCAAGGTGACCGCTGCC 2643  
Qy  
2836 TCAAGGTTTTCGCACTGCGCACTTCACCTGCTGTTGGTGGGCTGCTGATTCGAGAGGC 2895  
Db  
2644 TCAAGGTAGTCGCGACATGCGGCTGATGATGGTGGCGAGGACCTGACGCTGGCGGCT 2703  
Qy  
2896 TTTGCTCGGATCCACAAAGTACGACATCCAGATCTGTCATCGCTTCTCTCGCGCGC 2955  
Db  
2704 GTCGAAACCTGACCGTGAAGTGTCTTCCGGAATCTCGTGTGATCGATGCTGAAGGC 2763  
Qy  
2956 GAGCTTGTAACTTCAGGTGCTGGCGAGCGCACTGCGCACTCCGCGACCTGCGGACG 3015  
Db  
2764 GATCTCGGTCAATCGCGCGGCTGGCGAAGCGCTGCAAGAAAGCGCTGAAGGGT 2823  
Qy  
3016 CGCTCGAAGGCAAGCACCTCTGACGGAAGTTCTTGAGGA-----AG 3058  
Db  
2824 GAAGGCTTACAGGTTCTGCGGGTTCGTTCTGAGATGCCATCTCGATGCGGAA 2883  
Qy  
3059 AGCAGGCGCACTCGAGCTGATGATTCGAAGAACTGTCGAATAGCTCAACCGC---- 3114  
Db  
2884 CGGAAGGTGATCGAGCAAGCTGAGCGCAAGGTGATGATCTTCGATTTGCTGCTCTAT 2943  
Qy  
3115 CTGCTGTTCCCGAGCAACCGGAAGTTCCTCGACACGCTCGCGCTTCGCGCAACACC 3174  
Db  
2944 CTGATGATCCGAAGGTGTTCAACGATTTTCGCGCTGACCGCGAGACATACGCGCGGTC 3003  
Qy  
3175 TCTGCGTGGATGATCGTGAATTTCTTACGGCTGCTGTCGCTGGATGCTGATCTTGAATC 3234  
Db  
3004 TCCGTGCTGCCCACTGCTATTTCTATGGCATGGAGATGGCGAAGAG---CTGTTT 3060  
Qy  
3235 GCGCTGCAGATGTGCGCACTGCTGTTGCTGCTGGATGCTGATCTCTGAGCGAGAC 3294  
Db  
3061 GCGATATCGAGCGCGCAAGACGCTTGTCTCATGCTCAATCAGGCTTCGCTCCGCGATTGAC 3120  
Qy  
3295 GATAAGGCTATGCGCAATGTTGTGGCCAAAGTCAACGCGCGAGATCCGCCCAATTCGCTG 3354

Db 3121 GATAAGGCGATGGTGAACGGTGTCTTCGAGATCAAGCGCGCCGCGCATCAAGGTG 3180  
Qy 3355 CGTACCGCTCC---GTTGAGTCTGTCAACCGCAACCGCAGAAAGGAGAGATTCCTCCAAC 3411  
Db 3181 CCGGATCGCGCCCATGGCGCTTCGGCTCTGCGGTGCGCGCAAGGAGAGAAACCGGCAAT 3240  
Qy 3412 AAGGCGCATGTTGTCACACCATTCGCTGGTGTGTAC---CGTGACTGTGCTGAAGGT 3468  
Db 3241 CGGTCCCATATCGCGCGCGCGATGCCGTGGTGTCTATCAGCGCGCTTCATCAACCAAGGT 3300  
Qy 3469 GATGAGGTCAAGGCTGGAGATGCAAGTGGCAATCATCGAGGCTATGAAGATGGAAGCAACA 3528  
Db 3301 CAGAGGTCAAGCGCGCGAGCTGCTCTCTCGATCGAGCGCATGAAGATGGAAGCAAGCGG 3360  
Qy 3529 ATCACTGCTTCTGTGAGCGCAAAATCGATCGCGTGTGTGTCTCTGCTGCAACGAAGGTG 3588  
Db 3361 CTGATGCGCAACGAGCGGCAAGATCGCTGAAGTCTCTGCTGAAGCGCGGTGACCAAGATC 3420  
Qy 3589 GAAGTGGCGACTTGATCGTC 3609  
Db 3421 GACGCCAAGGATTTGTTGATC 3441

## RESULT 15

US-10-369-493-38504  
; Sequence 38504, Application US/10369493  
; Publication No. US20030233675A1

## GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 38504

; LENGTH: 3453

; TYPE: DNA

; ORGANISM: Agrobacterium tumefaciens

US-10-369-493-38504

Query Match 17.0%; Score 617; DB 15; Length 3453;  
Best Local Similarity 51.6%; Pred No. 1.2e-177;  
Matches 1776; Conservative 0; Mismatches 1590; Indels 75; Gaps 13;  
Qy 238 AAGATCTTGGTAGCAACCGCGGCGAAATCGCGTCCGTCCTTTCCGTCAGCACTCGAA 297  
Db 13 AAAATACCTTTGTCACACCGATCCGAAATGCGATCCGCTTTTCCCGGCGAGCAAGAG 72  
Qy 298 ACCGCTGACGACCGAGTATTTACCCCGTGAAGATCGGGATCATTCACCGCTCT 357  
Db 73 CTTGGGATCAAAACCGTTGCGATTTGGCGGAGAGGACAAAGTCTCTTTCCACCGCTTC 132  
Qy 358 TTGTGCTCTGAAGCTGTCCGCAATGGTACC-----GAAGGCTCAACA 399  
Db 133 AAGCGGATGATCTCTATCAGTCCGCGAGGCTCCGATCTCCGCAAGGATATGGGACCG 192  
Qy 400 GTCAGGCGTACTCGACATCGATGAATATTCGTCGAGCTAAAGATTAAGACAGAT 459  
Db 193 ATCGAGAGTATCTTTTCGATCAGAGGTTATCCCGGTGGCCAGCTCTCCGAGCGGAT 252  
Qy 460 GCATTTTACCCCGGATACGGCTCTCTGTCTGAATATGCCAGCTTCCCGCGAGTGTGCG 519  
Db 253 GCGATCCATCCCGTTATGCTCTCTGTCGAAAGCCCTGAGTTCGTGAGGCGCTGCAAC 312  
Qy 520 GAAAACGGCATTAATTTTATGCGCCCAACCCCGAGAGTCTTGATCTCACCGGTGATAAG 579

Db 313 AAGCGCGCATCACTTCATCGGCGCGCACCGGACAGATCGCCAGCTCGGCAACAAG 372  
Qy 580 TCTCGCGGTAACCGCGCGCAAGAGAGCTGGTCTGCGAGTTTGGCGGAATCCA---CC 636  
Db 373 GTGCGCGCGCGCAATCTGGCGATTTGGTGTGATGTCCCGTGTGTCCCGCGCACCAATCCT 432  
Qy 637 CCGAGCAAAAACATCGATGAGATGTTAAAGCGCTGAAGCGCTGAAGCGCAGACTTACCCCATCTTT 696  
Db 433 CTGCTGAGATATCGCGAAGTGGAGGTATGCCGAGGAATCGGTATTCCTCGCTCATG 492  
Qy 697 GTGAAGGCAAGTTCCCGTGGTGGGAGACGCGGTATGCGTTTGTGTTTCTACCTGATGAG 756  
Db 493 CTCGAAGGCTCTCTGGCGCGCGGTGGCGCGCATGCGCGCATTCGCAAAAAGGAAGAT 552  
Qy 757 CTTCCCAAAATAGCAACAGAGACATCTCTGAAGCTGAAGCGGCTTTTCGCGCATGGCGCG 816  
Db 553 CTCGCGCGGTGAGTGAAGGAGCGAAGCGCGGAGGCGCGCTTCGCGCAAGGACGAG 612  
Qy 817 GTATATGTGCAACGCTGTGATTAACCTTCAGCATATTTGAAGTGCAGATCTTTGGCGCAT 876  
Db 613 GTTTATCTGAAAAGCTGTCGAGCGCGCGCCCATGTTCGAGAGCGAGATTCCTCGCGCAT 672  
Qy 877 CACATCGAGAGATTTGTACACCTTTATGAAGTGTCTCTCACTGACGCTGCTCACCA 936  
Db 673 ACGCATGGCAATGTCTGTCATCTGTTCGAGCGGATTTGTTCGATCCAGAGCGCAATCA 732  
Qy 937 AAGATTGTGCAAAATTCGCGCACGACAGATTTGGATCCAGAACTCGGTGATCGCAATTTGT 996  
Db 733 AAGTCTGTCGAGCGCGCGCACCGTATCTGTCCGAAGCGCAGCGCGCAAGTCTGGCG 792  
Qy 997 GCGGATGAGTAAGTTCTGCGCTCCATTTGGTTACGAGCGCGCGGACCGTGGAAATC 1056  
Db 793 GCGTATTCCTGAAGATCGACGCGCGCAACATACATCGGCGCGCGCACTGTGCAATAT 852  
Qy 1057 TTGCTGCTAT---GAAAAGGCGCAACCACTCTTCATCGAAATGAACCCAGCTATCCAGGTT 1113  
Db 853 CTGATGATGCGGATACCGGCAATTTCTATTCAGAGTCAATCCGCTATCCAGGTG 912  
Qy 1114 GAGCACCGCTGACTGAAGATCTACCGAGGTGGACCTGGTGAAGCGCGCAGATGCGCTTG 1173  
Db 913 GAGCATACCGTTACCGAAGTGTGACCGCGCATCGATATCGTCAAGCGCGCAGATCCATAT 972  
Qy 1174 GCTGCTGCTGCAACCTT-----GAAGGAATTTGGTCTTGACC---CAAGATAAGATCAAG 1224  
Db 973 CTCGAAGAGAGCGGCGCATCGGACGCGGAGATCCGCGTGCCCAAGCAGAGAGATCCGC 1032  
Qy 1225 ACCACGCTGACGACTGCAAGTCCGCGCATCAACGAGAGATCCAAACAAACCGCTTCGCG 1284  
Db 1033 CTCACGCGCATGCGCTGCGATGCGCGCATCAGCAGGAGAGACCCGGAACACAAATTCAT 1092  
Qy 1285 CCAGATACCGGACTATCACCGGTACCGCTCACCGCGGAGCTGGCGTTCGCTTGAC 1344  
Db 1093 CCGGACTATGGCGGTATTAACCGCTATCGCTCGGCTTCGCTTCGCGATCCGCTGAT 1152  
Qy 1345 GGTG---CAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGTGTTGAAA 1401  
Db 1153 GGGGCGACGCTCTATACCGCGCGCGTCATCACCGCTTATTACGATCCGCTGCTGCTCAAG 1212  
Qy 1402 ATGACCTGCGGTGTTCCGACTTTGAACTGTGTGCTGCTGCAAGCGCGGTGGCT 1461  
Db 1213 GTCACGCGCTGGCGCGCGGAGCGCGGATGAAGCGATCAGCCGATGGATCGCGGTTCGCG 1272  
Qy 1462 GAGTTCACGCTGTCTGCTGTGCAACCAATTTGTTTCTTGGCTGCGTGTGCTGCGGAA 1521  
Db 1273 GAATTTGCTATCCGCGCGTTCGACCAACCTGACCTTCTCGAAGCATCATCGGTCAAC 1332  
Qy 1522 GAGGACTTCACTTCGAGGCGATCGCCACCGGATTCATTTGCGGATCACCGGACCTCTT 1581  
Db 1333 GACAGTTTCGCAACCAACACCTTATACGACGCGCTTCATCGATTCGACGCGCGAGTTGTC 1392  
Qy 1582 CAGGCTCCACTGCTGATGATGAGGAGGACGATCCTCGATTTACTTTGGCAGATGTCAAC 1641

Db 1393 GCGCAGGTCAGAGCGCGCCACCAAGCTTCTGACCTATCTGGCGGATGTGACC 1452  
Qy 1642 GTGAACAAGCCTCAAGGT-----GTGGCTCCAAAGAGATGTGACGTCTCTATCGATAAG 1695  
Db 1453 GTCAATGTCACCCAGAAACCAAGGTCGTGGAAGCCTGCGGACAAAGCGCGCAAGGCC 1512  
Qy 1696 CTGCGCTAACATCAAGGATCTGCATCTGCACCGCGTTCCTGAGACCGCTGAAGCAGCTT 1755  
Db 1513 ATCTGTCCTATATCGATGCGCGGACGCTGACGGCACCAAGCAATCTCTGGAAGAAGCTT 1572  
Qy 1756 GCGCCAGCGCGCTTCTGCTGCTGATCTCGCTGAGCAGAGCAGCACTGGCAGTCTACTGATACC 1815  
Db 1573 GCGCCTAAGGCTTGTGGGACTGGATGGCAACGAAAGCGCGTCTGTCACGATACG 1632  
Qy 1816 ACTTCCGGGATGCACACAGTCTTGTGCTTGGACCCGAGTCCGCTCAATTCGACATGAG 1875  
Db 1633 ACCATGCGGACGAGACACAGTCTGCTGTCGCGCACCCGCGTCCGACCATGATATGCCC 1692  
Qy 1876 CTTGGGCGAGAGCGCTGCAAGCTGACTCTCTGAGCTTTTTCCTGCGTGGAGGCTTGGGCG 1935  
Db 1693 CGGTTGCCAGGTTTATTCAGGCGCTTCGCGAGCTTCTGCTGGAATGCTGGGCG 1752  
Qy 1936 GCGCGACCTAGATGCGGATGCGTTCCTCTTCTTGGAGATCCGTTGGACAGGCTGAC 1995  
Db 1753 GGTGCGACCTTCGAGCTCTCAATGCGCTTCTCTGACCGAAGACCGCTGGGAGCGCTGTCA 1812  
Qy 1996 GAGCTGCGGAGCGGATGCGGAATGTAAACATTCAGATGCTGCTTCGCGCGCGCAACACC 2055  
Db 1813 CTATTCGCGAAGGTGCGCGGACCTGCTGTTGCGAGTCTTCTGCTGCTGCTGCGCAACGGC 1872  
Qy 2056 GTGGGATACACCCGCTACCGAGCTCGCTCGCGCGGCTTGTGTAAGAGCTGCCAGC 2115  
Db 1873 GTGCGTTACAAGAACTATCCGATATATGTCGTGAAGTATTTCTGCTCGTCAGCGCAAGG 1932  
Qy 2116 TCCGCGGTGGACATCTTCGCACTTCGACGCGCTTAACGAGCTCTCCAGATGCGTCCA 2175  
Db 1933 GCGGCGTGCATCTTTCGCTCTCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1992  
Qy 2176 GCAATCGACGAGCTCTGAGACCAACACCGCGGTAGCGAGGTGGCTATGGCTTATTTCT 2235  
Db 1993 TCGATGGATGCGATTCGCGGAGGAGACA---AGCTCTGCGAGGCAACCATCTGTACACC 2049  
Qy 2236 GGTGATCTCTGATCCAAATGAAGCTCTACCCCTGATTTACTTACTTAAAGATGGCA 2295  
Db 2050 GCGGATCTCTGATTCGCGAGTCTCGGAATATGATCTCAATATATACCAACCTCGG 2109  
Qy 2296 GAGGATCTCAAGTCTGCGCTCAATCTTGGCCATTAAGATATGGCTGTCTGTT 2355  
Db 2110 GTGAGCTGAAAGCGCGCGCCCATATCATCGCGTCAAGGATATGCGGCGCTGCTG 2169  
Qy 2356 CGCCGAGCTGCGTAACCAAGCTGCTACCGCACTGCGCGCTGAATTCGATCTGCCAGTG 2415  
Db 2170 AAACCGCGCGCGCCAGGTTCTGTTCAAGGCTTGGCGAGCGGACCGGCTGCGGATC 2229  
Qy 2416 CAGTGCACACCCAGCACTGCGGTGCGGAGCTGCGCAACCTACTTTGCTCAGCTCAA 2475  
Db 2230 CACTTCCACACCCATGATAGCTGCGGCAATTCGCGCGACCTTCTTGGCGCGCTGGAT 2289  
Qy 2476 GCTGCTGCAGATCTGTTGAGTGTCTTCGCACTCTCTGCGACCACTTCCAGCCCTCCAGCCA 2535  
Db 2290 GCGGCGGTGACGCGCTGATGCTGCGGATGATGCTTCTCGGCAATATGCTGCGAGCCC 2349  
Qy 2536 TCCCTGTCTGCCATTTGTGCTATTCGCGGACACCCGCTCGGATACCGGTTTGGAGCTC 2595  
Db 2350 TGCTCGGCTCGATGCTGGAGGCGCTCTCGGATCGGAACGGATACGGGTCTCGATACC 2409  
Qy 2596 GAGGCTGTTTCTGACCTCGAGCGGTACTGGAAGCAGTGGCGGACTGTACCTGCAATTT 2655  
Db 2410 GAATGGAATTCGCGCATTTCTCTTCTATGGAAGCGGTGCGCAATCAATATGCGGCGTTC 2469  
Qy 2656 GAGTCTGGAACCCAGCGCCCAACCGGTCTGCGCTTACCGGCAAGAAATCCAGGCGGACAG 2715  
Db 2470 GAGAGTGACCTCAAGGGCTCTGCTCGGAGTCTATCTGATGAATATGCGGCTGCGCAG 2529

Qy 2716 TTGTCACACTCGGTGTCACAGCGCACCGCACTGGGCTTGGCGATCGTTTGGAACTCATC 2775  
Db 2530 TTCAACCAACTCAAGGAACAGGCACGTTTCGTGGGTCTCGAAAGCGCTTGGCACGAGGTG 2589  
Qy 2776 GAAGCAACTTACGACGCGTTAATGAGATGCTGGGACGCGCCAAACAAAGCTCACCCCATCC 2835  
Db 2590 GCGCAGGCTATGCGGATGCTGACAGGATGTTTCGGGATATCGTCAAGGTGACGCGTCC 2649  
Qy 2836 TCCAAAGTGTGTTGGGACCTTCGCACTCCACCTGTTGTTGGTGGGTTGGATCCAGCAGAC 2895  
Db 2650 TCCAAAGTGTGTTGGGACATGCGCTGATGATGTTGGTGGGACGACCTGACGCGTGGCGAT 2709  
Qy 2896 TTTGCTGCGGATCCCAAAAAGTACGACATCCAGACTCTGTATCGCGCTTCTGCGCGCG 2955  
Db 2710 GTCGAAAACTCTGACGCTGAGTGTCTTCCCGGACTCGGTGTTATCGATGCTGAAGGCG 2769  
Qy 2956 GAGCTTTGTAACCTCTCAGGTGGTGGCGAGGCACTGCGGACCCGCGCATCGAAGGC 3015  
Db 2770 GATCTCGGTCAATGCGCGGCGGCTGCGCGAAGCGCTGCAAGAAAGCGCTGAAGGCT 2829  
Qy 3016 CGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCTCTGAGGA-----AG 3058  
Db 2830 GAAAAGCTTACAGGTTGTCGCGGTTCTCTGGAAGATGCCGATCTCGATGCCGAA 2889  
Qy 3059 AGCAGGCGCACTCTGACGCTGATGATTCCAAGAAAGCTGCAATAGCTTCAACCGC----- 3114  
Db 2890 CGAAAGGTGATCGAGACCAAGCTGAGCGCAAGGTGATGCTGAGTTTGCCTCTAT 2949  
Qy 3115 CTGCTGTTCCGGAAGCAACCGAAGAGTTCTCTGAGCACCGTCCGCGCTTCGCGCAACAC 3174  
Db 2950 CTGATGATCCGAGGCGCGGACAGCGTTGCTCATGCTCAATCAGGCTTCGTCGCGATTGAC 3126  
Qy 3175 TCTGCGCTGAGATGATCGTGAATTTCTACGCGCTGCTGCAAGGCGCGGAGACTTTTGATC 3234  
Db 3010 TCCGTGCTGCCCAACCCATGCTATTTCTATGSCATGAGGATGGCGAGAG---CTGTTT 3066  
Qy 3235 CGCTGCCAGATGTCGCGACCCCACTGCTTGTTCGCTGAGTGCATCTCTGAGCCAGAC 3294  
Db 3067 GCGGATATCGAGGCGCGGACAGCGTTGCTCATGCTCAATCAGGCTTCGTCGCGATTGAC 3126  
Qy 3295 GATAAGGATATGCGCAATGTTTGGCCAACTGTCGCAACCGCCAGATCCGCCCAATGCGTGTG 3354  
Db 3127 GATAAGGCGCATGTCGCGGTGTTCTTCGAGATCAACGCGCAGCGCGCGCATCAAGGTG 3186  
Qy 3355 CGTGACCGCTCC---GTTGAGTCTGTACCGCAACCGCAGAAAGGAGATTCCTCCAAC 3411  
Db 3187 CCGGATCGCGCCCATGCGCTTCGCGCTCTGCGGTGCGCGCAGGACAGACCCCGCAAT 3246  
Qy 3412 AAGGCGCATGTTGCTGCACCATTCGCTGCTGTTGTGTCAC---CGTGACTGTTGCTGAAGGT 3468  
Db 3247 GCGTCCCATATCGCGCGCGGATGCTGCTGTCATCAGCGCGCTTCTCATCAACGAGGT 3306  
Qy 3469 GATGAGTCAAGGCTGAGATGAGTTCGCAATCATCGAGCTATGAGATGGAGCAACA 3528  
Db 3307 CAGGAGTCAAGGCGCGGACGCTGCTCTGCTCGAGGCTATGAGATGGAAACGCGG 3366  
Qy 3529 ATCACTCTTCTGTTGACGCGCAAAATCGATTCGCGTGTGTTGTTCTCTCTGCAAGAGGTG 3588  
Db 3367 CTGATCCCGNACGCGACCGCAAGATCGTGAAGTTCTGTTGAAGCGCGGTGACCATC 3426  
Qy 3589 GAGGTGCGCACTTGTATCTGTC 3609  
Db 3427 GACGCCAAGGATTTGTTGATC 3447

Search completed: March 24, 2004, 06:08:42  
Job time : 1227 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 19:35:49 ; Search time 8670 Seconds  
(without alignments)  
12471.852 Million cell updates/sec

Title: US-10-045-072-1  
Perfect score: 3621  
Sequence: 1 tggggcgggtagatcctg.....tgaatgctgctgttcctaa 3621

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_hic.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_hic.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_man.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vri.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	609.8	16.8	3537	29 AY408287	AY408287 Mus muscu
2	606.2	16.7	3537	29 AY408285	AY408285 Homo sapi
3	442.4	12.2	3537	29 AY408286	AY408286 Pan trogl
4	432.2	11.9	2877	11 AK077665	AK077665 Mus muscu

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

c	5	240.6	6.6	888	29	CNS074FW	AL428754 clone BAO
c	6	236	6.5	903	13	BQ881271	BQ881271 AGENCOURT
c	7	223.4	6.2	859	28	AQ875053	AQ875053 V120F6 mt
	8	220.2	6.1	637	10	BF251052	BF251052 EST418309
	9	216.6	6.0	862	14	CA793824	CA793824 AGENCOURT
	10	216.4	6.0	675	9	AI109467	AI109467 GH08678.5
	11	214.8	5.9	744	12	BI103175	BI103175 602889415
	12	213.2	5.9	800	14	CK000115	CK000115 AGENCOURT
	13	211	5.8	640	12	BI618593	BI618593 RH49459.5
	14	208.2	5.7	715	12	BG910468	BG910468 602806142
	15	204.4	5.6	641	12	BJ063722	BJ063722 BJ063722
	16	202.8	5.6	635	14	CB576991	CB576991 AMGNNUC:C
	17	202.8	5.6	696	12	BP177875	BP177875 BP177875
	18	200.6	5.5	516	28	BZ293357	BZ293357 CG0285.v1
c	19	200.6	5.5	758	14	CF869970	CF869970 t1c021xb
	20	200.6	5.5	813	14	CB900127	CB900127 t1c021xb
	21	199	5.5	887	10	BF781506	BF781506 602104410
	22	198.4	5.5	546	14	CB279857	CB279857 t1c021xb
	23	196.6	5.4	515	9	AU202878	AU202878 AU202878
	24	195.6	5.4	508	12	BI621845	BI621845 RH53265.5
	25	193.6	5.3	665	12	BI605019	BI605019 RH70532.5
	26	192.8	5.3	2403	11	AK007782	AK007782 Mus muscu
	27	192.8	5.3	625	9	AI293242	AI293242 Mus muscu
	28	191	5.3	611	10	BE312170	BE312170 601152484
	29	190.6	5.3	915	13	BQ730755	BQ730755 AGENCOURT
	30	189.2	5.2	1030	29	CNS06M6L	AL405091 T3 end of
	31	188.4	5.2	949	13	BQ752438	BQ752438 EST633001
	32	188.2	5.2	921	13	BQ959497	BQ959497 AGENCOURT
	33	188	5.2	599	10	BF506511	BF506511 AT090003.5
	34	187.4	5.2	910	14	CD244370	CD244370 AGENCOURT
	35	185.2	5.1	910	14	CD244370	CD244370 AGENCOURT
	36	185	5.1	923	13	BQ734080	BQ734080 AGENCOURT
c	37	184.8	5.1	553	28	AQ501087	AQ501087 V27H5 mt
	38	184.4	5.1	895	10	BF535766	BF535766 602051271
	39	184	5.1	616	12	BJ124629	BJ124629 BJ124629
	40	183.8	5.1	813	12	BG761035	BG761035 602717545
	41	183.6	5.1	556	9	AI134619	AI134619 GH12134.5
	42	183.2	5.1	860	12	BI872878	BI872878 603397491
	43	182.4	5.0	725	12	BG599744	BG599744 EST504639
	44	181.6	5.0	609	9	AI526573	AI526573 u140b10.Y
	45	181.2	5.0	610	12	BI082302	BI082302 602877488

#### ALIGNMENTS

RESULT 1	AY408287	3537 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	Mus musculus PC gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
DEFINITION	AY408287				
ACCESSION	AY408287.1				
VERSION	GSS.				
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 3537)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferred nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PubMed	14671302				
REFERENCE	2 (bases 1 to 3537)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				

## JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

## FEATURES

source location/Qualifiers  
1..3537  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..>3537  
/gene="PC"  
/locus\_tag="HCM3173"

## gene

## ORIGIN

Query Match 16.8%; Score 609.8; DB 29; Length 3537;

Best Local Similarity 48.5%; Pred. No. 2.2e-131;

Matches 1361; Conservative 0; Mismatches 1420; Indels 24; Gaps 6;

QY 233 TCAGAAAGATCTGTAGCAACCGCGGAAATCGCGTCCGCTTCGTCGACGAC 292  
DB |||||  
QY 110 TCAGAAAGATCTGTAGCAACCGCGGAAATCGCGTCCGCTTCGTCGACGAC 169  
DB |||||  
QY 293 TCAGAAACCGGTGACGACGAGTGTATTTACCCCGTGAAGATCGGGATCATTCAC 352  
DB |||||  
QY 170 CAGAGCTGGGTATCCGACAGTGGCTGTCTACTCGGACGACACACAGGCGAGATGCACA 229  
DB |||||  
QY 353 GCTCTTTTGTCTCTGAGCTGTCCGATTTGTTACCGAGGCTCCAGCTCAAGCGGTACC 412  
DB |||||  
QY 230 GGCAGAAAGCTGATGAAGCTTACCTTATTTGGCGGTGCGCTGTGAGGCTTACC 289  
DB |||||  
QY 413 TGGACATCATGAATATTCGGTGCAGCTTAAAGATTTAAAGACGATGCCATTTACCCGG 472  
DB |||||  
QY 290 TGCACATTCACGACATCATCAAGTGTGCGCAAGAAATGGTGTAGATCGGTGCATCTG 349  
DB |||||  
QY 473 GATACGGCTTCTGTCTGAAATGCCAGCTTGCCTGCGGAGTGTGCGGAAACGCGATTA 532  
DB |||||  
QY 350 GCTATGGGTCTCTCAGAGCGAGCAGACTTTGSCCAGGCTTGCAGATGCTGAGTCC 409  
DB |||||  
QY 533 CTTTATTTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATTAAGTCTGCGCGGTAA 592  
DB |||||  
QY 410 GGTTCATTTGGTCCAGCCAGAGGTTGTCGCGAAGTGGAGACAGTGGAGCCCGGG 469  
DB |||||  
QY 593 CCGCCCGGAAGAAGGTGTGTCTGCGAGTTTGGCGGATC---CACCCGAGCAAAACA 649  
DB |||||  
QY 470 CCAATGCCATCGCTGAGCGGTTCAGTGTGTCGCGGACGAGCTCCGCCATCAGCTCCC 529  
DB |||||  
QY 650 TCGATGAGATCGTTAAAGCGGTGAAGGCGACAGCTTACCCCATCTTTGTGAAGGCAATTG 709  
DB |||||  
QY 530 TGCACAGGGCGCATGATGTTCTCAACACCTACGGCTTCCCTATTATCTTCAAGGCGGCT 589  
DB |||||  
QY 710 CCGGTGTGCGGACCGGATATGCGTTTGTGTTTCACTGATGAGCTTGCAGATTTAG 769  
DB |||||  
QY 590 ACGAGGTGGGGCCCGGCGATCGCGGTCTGTCATAGCTATGAGGAGTTGGAAGAAATT 649  
DB |||||  
QY 770 CAACAGAGCATCTCTGTAAGCTGAAGCGCTTTCGCGCATGCGCGGTATATGTCGAAC 829  
DB |||||  
QY 650 ACACCGGCGCTACTCCGAGGCTTCCGAGGCTTTCGGAATGAGGCTTGTGTTGAGAA 709  
DB |||||  
QY 830 GTGCTGTGATTAACCTTCAGCATATTGAAGTGAGATCTTGGCGATCATCTGAGAGAG 889  
DB |||||  
QY 710 AGTTCAATTGAGAAGCCAAAGSCAATTCAGGTGCGATCTCTAGGGACCAAGTATGGGAACA 769  
DB |||||  
QY 890 TTGTACACCTTTATGAAGCTGACTGCTCACTGCGAGCTGCTCACCAAAAGTTGTCGAAA 949  
DB |||||  
QY 770 TCTTGCACTGTACGAGCGAGATCTGTCATTCAGGCTTGGCACCGAAGAGTGTGATAGA 829  
DB |||||  
QY 950 TTGCGCGCAGCAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCGAGTAA 1009  
DB |||||  
QY 830 TCGCCCTGCTACCCACCTGGATCCCAACTTCGCTCAGCTCTCACCAGTGCATCTGTCA 889  
DB |||||  
QY 1010 AGTCTGCGCCTCCATTTGTTACGAGGCGGGAACCGTGAATCTTGTGTCATGAAA 1069  
DB |||||  
QY 890 AACTTGCAGAGGAGTAGGCTATGAGAAACCGCGGCACTGTGGAGTTCTCTGGTGGACAGC 949  
DB |||||

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

1070 AGGGCAACCAACGCTCTTTCATCGAAATGAACCCACGATCCAGGTTGAGCACACCGTGACTG 1129

950 ACGCAAGCACTACTTTCATCGAGGTCAATTCGCGCTGAGTGGAGCACACGTCACCG 1009

1130 AAGAGTCAACCGAGTGGACCTGTTGAAGGGCGAGATGGCTTGGCTGCTGCTGCAACT 1189

1010 AGGAGATCAAGANN 1069

1190 TGAAGGAATTGGGTCTGACCCCAAGATAAGATCAAGACCCACGTTGACAGCTGCACTGCC 1249

1070 NNN 1129

1250 GCATCACACGGAAGATCCAAACAACGGCTTCGGCCAGATACCGGAATATCACCGGCT 1309

1130 NNN 1189

1310 ACCGCTCACACGCGAGCTGGCGTTTCGTCTTGAACGTTGACGTCAGCTCAGTTCGTCGCGAAA 1369

1190 TCCGAGTGTGAGGCGCATCCGCTTGACAAGCCCTCTGCAATTCAGGCGCGCTG 1249

1370 TCA---CCGACACATTGTGACTTCATGCTGTTGAAATGACCTGCGCTGCTCCGACTTTG 1426

1250 TCATATCGCCCACTATGACTCTCTGCTGCTCAAGTCAATTCGACACGCGAAAGACCCACC 1309

1427 AAATCTGCTGTGCTCGTGACACAGCGCGCTTGGCTGAGTTTCAACGCTGTCTGTTGCAA 1486

1310 CCACAGCTGCCACCAAGATGAGCAGAGCCCTGCGGAGTTCGCTGTCGAGGTGTAAAGN 1369

1487 CCAACATTTGGTTTCTGCTGCTGCTGCTGCGGGAAGAGACTTCACTTCCAAGCGCATCG 1546

1370 NNN 1429

1547 CCACCGGATTCATTGCGGATCACCCGCACTCTTCAAGGCTCCACCTGCTGATGATGAGC 1606

1430 NNN 1489

1607 AGGACGCACTCTGGAATTAATTGGCAGATGTCAACGTTGAACAGCTCATGTTGCTGCTC 1666

1490 NNN 1549

1667 CAAAGGATTTGCGAGTCTCTATCGATAAGCTGCTAACTCAAGGATC-----TGCCAC 1720

1550 CTGTCAATGTGAGCCCGAGCTCTGTGATCTCTGCTGCTCTCTGCTGCTGCTGCTGCTGCT 1609

1721 TGCCACGCGTTCCCGTGACCGCTGAAAGAGCTTGGCCGAGCGCGCTTGTGCTGCTGATC 1780

1610 CTCAGCTGTTTCAGGACATCTTCTGCGAGAGGCGCAGAGGCTTTGCGCGAGCTG 1669

1781 TCCGTGAGCAGGACGCACTGGCAGTTTACTGATACCACTTTCGCGGATGCAACCACTCTT 1840

1670 TGGGAATACCAAGGCTCTCTGTGATGACCAACTTTCGCGGATGCCACCACTCTC 1729

1841 TGCTTGCACCCGAGTCCGCTCAATTCGCACTGAAGCTGCGGCGAGAGCGCTGCGAAGC 1900

1730 TACTGGCCACTAGAGTGGCACAATGATCTCAAAAAGATTGCGCCCTATGTTGCGCCACA 1789

1901 TGACTCTGAGCTTTGCTCGTGAGGCTTGGCGGCGGAGCTTACGATGTGGCGATGC 1960

1790 ACTTCAACAAGCTCTTCAAGCATGAGAACTGGGAGGCGGCACTTTCGAGCTTGCATGCT 1849

1961 GTTTCTCTTTTGAAGATCCGTTGGGACAGGCTCGACGAGCTGCGGAGGCGGATGCGGAATG 2020

1850 GCTTCTGTACGAGTGGCTTGGCGGCGGCTCCAGGAGCTCCGCGAGCTTATCCGGAACA 1909

2021 TAAACATTCAGATGCTCTTTCGCGGCGGCAACCGTGGGATACCCCGTACCCAGCT 2080

1910 TCCGCTTCCAGATGCTACTGAGGCGGCGCAATGCTGTGGGCTTACCACTACCTGACA 1969

2081 CCGTCTGCGCGCTTTGTTTAAAGGAAGCTGCCAGCTCCGCGCTGGAATCTTCCGCTCT 2140

1970 ACGTGTCTTCAAGTTCTGTGAGGTGGCCAAAGAGATGATGATGAGCTCTTCCGAGTCT 2029

DB

Qy	2141	TCGACGGCGTTAACGAGGCTCCCGATGCGGTCCAGCAATCGACGCGAGTCTCTGGAGACCA	2200
Db	2030	TTGACTCCCTTCAACTACTTTGCCAAACATGCTGTGGGCATGGGAAGCAG---	2086
Qy	2201	ACACCGCGGTAGCCGAGGTGGCTATGGCTTATTTCTGTGTGATCTCTCTGATCCAAATGAAA	2260
Db	2087	CTGCGGGTGTGGTGGAGGCTGCCATCTCATACACGGGGACGTGGTGACCCCTAGTGGCA	2146
Qy	2261	AGCTCTACACCTTGGATTACTTAAAGATGACAGAGGATGCTCAAGTCTGGCGCTC	2320
Db	2147	CTAAATACTCACCTGGAGTACTCATATGGGCTTAGCTGAAGAACTGGTGGAGCTTGGC	2206
Qy	2321	ACATCTTTGGCCATTAAAGATATGGCTGGTCTGTTCCGCCAGCTGCGGTAAACCAAGCTGS	2380
Db	2207	ACATCTCTGCAITTAAGGACATGGCGGCCCTGCTGAAGCCCTGCCGCTGCACCATGCTGG	2266
Qy	2381	TCAACCGACTGCGCCGCTGAAT---CGATCTGCCAGTGCACGTCGACACCCACGACACATG	2437
Db	2267	TCAGCTCCCTCCGGAGCCGATTTCCCGACCTCCCATCTGCACATCATCCCATGATACAT	2326
Qy	2438	CGGCTGCCAGCTGGCAACCTACTTTTGCTGCGAGCTCAAGCTGGTGGCAGATGCTGTTGACG	2497
Db	2327	CAGGGCAGGTGTGGCAGCCATGTGCGCCTGTGCAAGCAGGGGCTGATGTTGTGGACG	2386
Qy	2498	GTGCTTCCGCACACACTGCTGTGGCACACCTCCACGCCATCCCTGTCTGCCATTTGTTGCTG	2557
Db	2387	TGGCAGTAGACTCCATGTCTGGGATGACTTCACAGCCAAAGCATGGGGCCCTGTGSCCT	2446
Qy	2558	CATTGCGGCAACCCGTGGGATACCGGTTTGAGCCTTCGAGGCTGTGTTCTTGACCTCGAGC	2617
Db	2447	GTACCAAGAGGACTCTTTTGACACAGAGTACCCCTTGGAGCGGTGTTTGACTACAGTG	2506
Qy	2618	CGTACTGGGAAGCAGTGGCGGACTGTACCTGCCATTGTA-----GTCTGGAAACCCCG	2671
Db	2507	AGTACTGGGAAGGGGCTCGGGACTGTATCGCAGCTTCGATTCGACGGCTACCATGAAGT	2566
Qy	2672	GCCCAACGGTGGGTCTACCGCCACGAATCCAGCGGACAGTGTCTCCAACTGCGTG	2731
Db	2567	CTGCAACTCCGACCTGTATGAGAAATGAGATTCCAGGGGGCCAGTACACCAACTGCACT	2626
Qy	2732	CACAGGCCACCGCACTGGGCGCTTCGGATGGTTTCGAACCTCATCGAAGACAACTACGACG	2791
Db	2627	TCCAGGCCCATAGCATGGGCTTGGCTCCAGTTTCAAGGAGGTCAAGAGGCTTATGTGG	2686
Qy	2792	CCGTTAATGAGATGCTGGGACGCCCAACCAAGGTACCCCATCTCTCAAGGTTGTGGCG	2851
Db	2687	AGGCTAACCAAGATGCTGGGGACCTCTCATCAAGGTGACACCATCTCTCAAGATTGGGGG	2746
Qy	2852	ACCTTCGCACCTCCACTGGTTGGTGGGCTGGATCCAGCAGACTTTGTCTGCCGATCCAC	2911
Db	2747	ACCTGGCCCAAGTTTCATGTGTGAGAATGGGTTTGAGCCGGGCAGAGGCTCAGGCAG	2806
Qy	2912	AAAAGTAGCATCCCGACATCTCTCATCTCGGTTTCTTGGCGGGCGAGCTGGTAACCTC	2971
Db	2807	AAGAGCTGTCTTCCCCCGCTCTGTGGTGGAGTTTCTGTGAGGGCTCATATTGGCATCCCC	2866
Qy	2972	CAGGTGCTGCCAGAGCACTGCGCACCCCGCACTGGGAAGGCC	3016
Db	2867	ATGGGGTTTCCCTGAGCGCTTTCCGCTCTAAGGTGCTAAAGGAC	2911

RESULT 2	AY408285	3537 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	AY408285				
DEFINITION	Homo sapiens PC gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
ACCESSION	AY408285				
VERSION	AY408285.1	GI:39764256			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				

REFERENCE	1 (bases 1 to 3537)
AUTHORS	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarimal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL	Science 302 (5652), 1960-1963 (2003)
PUBMED	14671302
REFERENCE	2 (bases 1 to 3537)
AUTHORS	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarimal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1..3537
gene	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>3537 /gene="PC" /locus_tag="HCM3173"
ORIGIN	
Query Match	16.7%; Score 606.2; DB 29; Length 3537;
Best Local Similarity	48.7%; Pred. No. 1.5e-130;
Matches 1365; Conservative	0; Mismatches 1416; Indels 24; Gaps 7
QY	233 TCAAAAAGATCTTGGTAGCAACCGCGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCAC 292
DB	110 TCAAGAAAGTCATCGTGCACCAAGAGGTGAGATTGCCATCCGTGTGTTCGGGCGCTGCA 169
QY	293 TCGAAACCGGTGCAGCCAGCGTAGCTATTATACCCCGGTGAAGATCGGGGATCATTTCCACC 352
DB	170 CGGAGCTGGGCGATCCGCACCGTAGACCACTACTCTTGAGCAGGACACGGGCCAGATGCACC 229
QY	353 GCTCTTTTGTCTTGAAGCTGTCGCGCATTTGGTACCGAAGGTCTACCAAGTCAAGGGGTACC 412
DB	230 GGCAGAAAGCAGATGAAGCCCTATCTCATCGCGCGCGGCTTCGCCCGCTGCGAGGCTTACC 289
QY	413 TGGACATCGATGAATTTATCGTGCAGCTAAAGTTAAAGCAGATGCCATTTATCCCGG 472
DB	290 TGCATATCCAGACATCATCAAGGTGGCCAGGAGAACACGTAGATCAGTGCACCTTG 349
QY	473 GATACGGCTTCCTCTCTGAAATGCCAGTTTGCCTCGCGAGTGTGCGGAAAACGCGATT 532
DB	350 GCTACGGGTTCCTCTCTGAGCGAGCGGACTTCGCCCGCTGCCAGGATGCGAGGGTCC 409
QY	533 CTTTATTTGGCCCAACCCAGAGTTCTTGATCTCACCGGTGATAAGTCTCGCGGGTAA 592
DB	410 GGTATTATGGCCCAAGCCCAAGATGGTTCGCGAAGTGGGAGACAAGTGGAGGCCCGGG 469
QY	593 CCGCGCCGCAAGAAGCGCTGGTCTGCCAGTTTGGCGG---AATCCACCCCGAGCAAAAACA 649
DB	470 CCATCGCCATTGCTCGGGTGTTCCTCGTGTCCCTTGGCACAGATGCCCCCATCATCGTCC 529
QY	650 TCGATGAGATGTTAAAGCCCTGAGGCCAGACTTACCCATCTTTGTGAGGCGAGTGTG 709
DB	530 TGCATGAGGCCCAAGGTTCTCCAAACACTACGGCTTCCCATCATCTTTCAAGGCGGCT 589
QY	710 CCGGTGGTGGCGGACCGCGTATGGTTTTGTGTTCACCTGATGAGCTTCGCAAAATTAG 769
DB	590 ATGGGGGTGGAGGGCGTGGCATAGGGTGGTGCACAGCTACGAGGAGTGGAGGAGAATT 649
QY	770 CAACAGAAAGCATCTCGTGAAGCTGAGCGGCTTTCCGCGAGTGGCGCGGTATATGTGCAAC 829
DB	650 ACACCCGGGCGCTACTCAGAGGCTCTGGCGGCTTTGGGAATGGGGCGCTGTGTGTGGAGA 709

QY 830 GTGCTGTGATTAACCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACAAGTGCAGAG 889  
DB 710 AGTTTCATCGAGAGCCACGGCAGATCGAGGTGCAGATCTTGGGGAGCCAGTATGGGAACA 769  
QY 890 TTGTACACCTTTATGAAGTGAAGTCTCACTGCAGCGTCTGTCACCAAAAGTCTCGAAA 949  
DB 770 TCCCTGCACTGTACGAGCGAGACTGCTCCATCCAGCGCGGCAACAGAGGTGTTCGAGA 829  
QY 950 TTGGCGCAGCAGCATTTTGGATTCAGAACTTCGCGTATCGCATTTTGTGCGGATCGAGTAA 1009  
DB 830 TTGGCCCCCGCCACCTGGAACCGCGAGCTTCGGACTCGGCTCACAGCGACTCTGTGA 889  
QY 1010 AGTTCTGCGCGTCCATTTGTTTACAGGCGCGGGAACCGTGGAAATCTTGGTCCATGAAA 1069  
DB 890 AACTCGCTAAAAGTGGGCTACGAGAACGAGGCAACCGTGGAGTCTCTGTGTGACAGGC 949  
QY 1070 AGGGCAACCACTCTTCATCGAAATGAACCCAGTATCCAGGTTAGACACCGTGAATG 1129  
DB 950 ACGGCAAGCACTACTTTCATCGAGGTCAAATCCCGCTTCAGGTGAGACACGCTCACAG 1009  
QY 1130 AAGAAGTCAACCGAGGTGGACCTGGTGAAGCGCAGATCGCTTGGCTGCTGTGCAACCT 1189  
DB 1010 AGGAGTCAACCGANNN 1069  
QY 1190 TGAAGGAATTTGGGTCTGACCCCAAGATAAGATCAAGACCCAGCGTGACGACTGCGC 1249  
DB 1070 NNN 1129  
QY 1250 GCATCACACGGAAGATCAAACAAGCGCTTCCGCCAGATACCGGAATATCACCGCGT 1309  
DB 1130 NNN 1189  
QY 1310 ACCGCTCACAGCGGAGCTGGCTGTGCT---TGACGGTGCAGCTCAGCTCGGTGGCG 1366  
DB 1190 TCGGAGCGAGAGGAGCATCGGCTGGATAATGCTTCCGCTTCCAAAGGAGCGC 1249  
QY 1367 AATACACCGCACATTTGATCCATGCTGCTGTAATAATGACCTCGCGTTCGCACTTG 1426  
DB 1250 TCACTCGGCCCACTACGACTCCCTGCTGTCAAAGTCAATTTGCCCAAGGACACACC 1309  
QY 1427 AAATGCTCTGCTGCTGCACAGCGCGCTTGGCTGAGTTCAACCGTGTCTGTGTGCAA 1486  
DB 1310 CCAGCGCGCCCAAGATGAGCAGGCGCTTTCGGAGTTCCGCTCCGAGGTGTGAAGN 1369  
QY 1487 CCAAATTTGGTCTTTCGCTGCTGCGGGAAGAGACTTCACTTCCAGCGCATCG 1546  
DB 1370 NNN 1429  
QY 1547 CCACCGAATTCATTTCCGATCACCGCATCTCTTCAGGCTCCACTGCTGATGAGC 1606  
DB 1430 NNN 1489  
QY 1607 AGGAGCGCATCTGGATTAATTGGCGAGATGTACCGTGAA-----CAAGCTCATGGTGTG 1662  
DB 1490 NNN 1549  
QY 1663 CGTCCAAAGAGTGTTCAGCTCCTATTCGATAAGCTG--CCTAAATCAAGGATCTGCCAC 1720  
DB 1550 CCGTCAAGCGCAGCCCAAGCCCAAGCCCGCTTGTCCCTGCTGAGTCCCATAGGCCGCG 1609  
QY 1721 TGCCAGCGGGTCCCTGTGACCGCTGAAGCAGGTTGGCCAGCGGGTGTGCTGTGATC 1780  
DB 1610 CCGCGCTGGTTTTCAGAGATCTCTGCTCGAGAGGGGCTTGGGGCTTGTCTGAGCTG 1669  
QY 1781 TCCGTGAGCAGGACGCACTGGCAGTACTGATACCACTTCCGCGATGCACACCACTT 1840  
DB 1670 TGGGGAACCAACCGGGCTGCTGCTGATGGAACAGACTTTCAGGAGCGCCACAGTCA 1729  
QY 1841 TGCTTGCAACCGAGTTCGCTCATTTGCACTGAGGCTTCGGGAGAGGCGCTGCAAGC 1900  
DB 1730 TGTGGGCCACTGCTGTGCGCACCCAGATCTCAAAAGATCGCCCTATGTTGCCACA 1789  
QY 1901 TGACTCCTGAGCTTTTGTTCGCTGGAGCGCTGGGCGCGGACCTACGATGTGGCGATGC 1960

DB 1790 ACTTCAGCAAGCTCTTTCAGCATGGAGAACTGGGAGAGGCCACGTTTGACGTGCGCATGC 1849  
QY 1961 GTTTCCTCTTTTTCAGGATCCGTGGGACAGCTCGACGAGTGCAGGCGGATGCGGAATG 2020  
DB 1850 GCTTCGTATGATGTCCTCGCGCGGCTCGAGAGTTCGGGAGCTCATCCCAACA 1909  
QY 2021 TAAACATTCAGATGCTTTCGCGCGGCAACACCGTGGGATACACCGGTACCCAGACT 2080  
DB 1910 TCCCTTTTCAGATGCTGCTCGGGGGCCATGCTGTGGGCTACACCACTACCCAGACA 1969  
QY 2081 CGCTTCGCGCGCTTTTAAAGGAAGTGCAGCTTCGCGGTGACATCTTTCGCAATC 2140  
DB 1970 ACGTGTCTTCAAGTTCTGTGAAGTGGCCAAAGAGATGGCATGATGTCTTCCGTGTG 2029  
QY 2141 TCGACGGCTTAAAGCGTCTCCAGATGCGTCCAGCAATCGACGAGTCTCTGGAGACA 2200  
DB 2030 TTGACTCCCTCAACTTTCGCCAATGCTGCTGGGCAATGGAGCGG---CAGGAAGTG 2086  
QY 2201 ACACCGCGGTAGCGAGGTGGCTATGGCTTATTTCTGTGGTGAATCTCTGATCAAAATGAAA 2260  
DB 2087 CCGAGGCGGTGTGGAGGCTGCATCTCATACAGCGCGACGCTGGCCGACCCAGCCGCA 2146  
QY 2261 AGCTCTACACCGCTGATTAATTAACCTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGTC 2320  
DB 2147 CCAAGTACTCACTGCAAGTACTATATGGGCTTGGCGAGAGCTGTGCGAGTGTGCACCC 2206  
QY 2321 ACATCTTGCCCATTAAGGATATGGTGTCTCTTGGCCAGCTGCGGTAAACAGCTGG 2380  
DB 2207 ACATCTGTGCTCAAGGACATGGCGGCTGCTGAAGCCACGCGCTGCACCATGCTGG 2266  
QY 2381 TGACCGCATGCG---CGGTGAATTCGATCTGCGAGTGCAGCTGCACACCCAGCACATG 2437  
DB 2267 TCAGCTCCCTCGGAGACCGCTTCCCGACTTCCCTGAGCATCCACCCAGCACAT 2326  
QY 2438 CGGCTGGCGCAGCTGGCAACTTCTGCTGCAGCTCAAGCTGGTGCAGATGCTTGTGACG 2497  
DB 2327 CAGGGCAGCGGTGGCAGCATCTGCTGCTGCGCCAGCTGGAGTGTGTTGGATG 2386  
QY 2498 GTGCTTCGCAACACTGTCTGGCAACACTCCAGGCATCCCTGTCTGCCATTTGTTGCTG 2557  
DB 2387 TGGCAGCTGATTCATGCTGGGATGACTTACAGCCAGCATGGGGGCGCTTGGTGGCT 2446  
QY 2558 CATTCGCGCACACCCCTCGCATACCGGTTTTCGAGCTCGAGGCTGTTCTGACCTCGAGC 2617  
DB 2447 GTACAGAGGACTCCCTTGGACAGAGGTGCCATGGAGCGCTGTTGACTACAGTG 2506  
QY 2618 CGTACTGGGAAGCAGTGGCGGACTGTACTGTCATTTGAGTCTGGAAACCCAGAGCCCAA 2677  
DB 2507 AGTACTGGGAGAGGCGCTCGGGGACTGTACGCGGCTTTCGACTGCACGCGCCCATGAAGT 2566  
QY 2678 CCGG-----TCGCGCTACCGCCACGAAATCCAGGCGGACAGTTGTCCAACTCGCTG 2731  
DB 2567 CTGGCAACTCGGACGTGTATGAATAATGAGATCCAGGGGCGCAGTACACCACTGCACT 2626  
QY 2732 CACAGGCCACCGCACTGGGCTTTCGAGTCTGGTTCGAACTCATTCGAGAGCACTAGCGAG 2791  
DB 2627 TCCAGGCCACAGCATGGGGCTTGGCTCCAAAGTTCAAGGAGGTCAAGAGGCGCTATGTG 2686  
QY 2792 CGTTAATGAGATGCTGGGACGCGCCCAAGAGTCCCACTCCCTCCAAAGTTGTTGGCG 2851  
DB 2687 AGGCCAACAGATGCTGGGAGTCTCATCAAGTGAACGCTCTTCCAAAGTCTGGGGG 2746  
QY 2852 ACCTGCACTCCACTGTTGTGGTGGGCTGTGATCCAGCAGACTTTGTGCGGATCCAC 2911  
DB 2747 ACCTGGCCCACTTATGTTGCAAGATGGATTCAGCGCGGCGAGAGGCTCAGGCGG 2806  
QY 2912 AAGATACAGCATCCAGACTCTGTCATCGGTTCTGCGGCGGAGCTTGGTAACCTTC 2971  
DB 2807 AAGAGTGTCTTTTCCCGCTCCGTGGTGAAGTCTCTGAGGCTTACATCGGTTGCTCCCC 2866  
QY 2972 CAGGTGCTGGCAGAGCCACTCGGCACCGCGCACTGGAAGGCC 3016



Db 2867 ATGGGGGGTTCGCCGAACCCCTTTCGCTCTAAGTACTGAAGGACC 2911

RESULT 3  
LOCUS AY408286 3537 bp DNA linear GSS 15-DEC-2003  
DEFINITION Pan troglodytes PC gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY408286  
VERSION AY408286.1 GI:39764257  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Mammalia; Primates; Catarrhini; Hominidae; Pan.  
REFERENCE 1 (bases 1 to 3537)  
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B.,  
Perrera, S., Wang, G., Zheng, X.H., White, T.J., Sainsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 3537)  
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B.,  
Perrera, S., Wang, G., Zheng, X.H., White, T.J., Sainsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source Location/Qualifiers  
1..3537  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>3537  
/gene="PC"  
/locus\_tag="HCM3173"

gene

ORIGIN

Query Match 12.2%; Score 442.4; DB 29; Length 3537;  
Best Local Similarity 41.5%; Pred. No. 4e-92;  
Matches 1004; Conservative 0; Mismatches 1400; Indels 18; Gaps 6;

Qy 233 TCAGAAAGATCTTGGTAGCAACCGCGCGAATCGCGTCCGCTTTCGCTGAGCAC 292  
Db 110 TCAAGANNCTATGGTGGCAACAGAGGTGAGATGCCATCCGTGTGTTCCGGGCTGCA 169  
Qy 293 TCAGAAACCGGTGCAGCCACCGTAGCTATTATCCCGCTGAGATCGGGGATCATTCAC 352  
Db 170 CGGAGCTGGGCATCCGACCGTAGCCATCTCTGAGAGGACACGGGCCAGATGCACC 229  
Qy 353 GCTCTTTTGTCTGAGAGCTGTCGGATTTGGTACCGAAGCTCACCAGTCAAGGCTACC 412  
Db 230 GGCAGAAAGCAGATGAAGCCCTATCTCATCGCGCGCGCCCTGGCCCCCGTGCAGGCTTACC 289  
Qy 413 TGACATCGATGAATATCGTGCAGTAAAGATTAAAGCAGATGCCATTTACCCGG 472  
Db 290 TGCACATCCAGACATCATCAAGTGGCCAGGAGACACGTAGATGCAGTACACCTTG 349  
Qy 473 GATACGGTCTCTCTGTAATGCCAGCTTGCCTCCGAGTGTGCGGAAACCGGATT 532  
Db 350 GCTACGGGTCTCTCTGAGCGAGCGACTTCGCCCGAGGCTGCCAGNATGCAAGGGTCC 409  
Qy 533 CTCTTTATTTGGCCCAACCCAGAGTCTTGTATCTCACCGGTGATAAGTCTCGCGCGGTAA 592  
Db 410 GGTATTTATGGCCCAAGCCAGATGGTCCGACAGATGGGAGACAGGTGGAGGCCCGGG 459  
Qy 593 CCGCGCGAGAGAGGCTGTGTCGACGTTTGGCGG---AATCCACCCCGAGCAAAACA 649

470 CCATCGCCATTTGTCGGGGTGTTCCTCGTTCCTGCGCAGATGCCCATCACGTCCC 529  
650 TCGATGAGATCGTTTAAAGCGCTGAAGCCAGACTTACCCCATCTTTGTGAAGCAGTTG 709  
530 TGCATGAGGCCCCAGAGTCTTCCAAACACTTACGGCTTCCCATCATCTTCAAGGCGCCT 589  
710 CCGTGTGTGGGAGCGCGGTATGCGTTTGTGCTGCTTCACTGATGAGCTTCGCAATAG 769  
590 ATGGGGGTGGAGGCGGTGGCATGAGGGTGTGTCACAGCTACGAGNNNNNNNNNNNNNN 649  
770 CAACAGAAAGCATCTCGTGAAGCTGAAGCGCTTTCGGCGATGCGCGGTATATGTGAAAC 829  
650 NNN 709  
830 GTGCTGTGATTAAACCTCAGCATATTGAAGTGAGATCTTGGCGGATCACACTGGAGAAG 889  
710 NNN 769  
890 TTGTACACCTTTATGAACGTGACTCTCCTGACGCGCTGTCACCAAAAAGTTGTGAAA 949  
770 NNN 829  
950 TTGCGCCAGCAGCATTTGGATCCAGAACTGGGTGATCGCATTTGTGCGGATGCGATAA 1009  
830 TTGCCCCCGCGCCNN 889  
1010 AGTTCTCGCGTCCATTTGTTTACAGCGCGCGGAAACCGTGGATCTTCTGTCGATGAA 1069  
890 NNN 949  
1070 AGGCAACCACTTTCATCGAAATGAACCACTGATTCAGAGTTGAGCACACCGTACTG 1129  
950 ACGCAAGCACTACTTTCATCGAGGTCACTCCCGCTGACAGTGGAGTGGAGTGGAGTGG 1009  
1130 AAGAGTCAACCGAGTGGACCTGCTGAGGCGCAGATGCGCTTGGCTGCTGCTGCAACCT 1189  
1010 AGGAGTCAACCGAGTGGACCTGCTGAGGCGCAGATGCGCTTGGCTGCTGCTGCAACCT 1069  
1190 TGAAGGAATTTGGTCTGACCAAGATAAGATCAAGACCCAGCTGAGCATGCGATGCC 1249  
1070 NNN 1129  
1250 GCATCACACGGAAGATCCAAACAAACCGCTTCCGCCAGATACCGGAACATATCACCGCT 1309  
1130 NNN 1189  
1310 ACCGCTCACAGCGGAGCTGGCGTTCTGCTCT---TGACGGTGCAGCTCAGCTCGGTGGCG 1366  
1190 NNGGAGTGGAGAGGGCATGGGCATNNGCCTGGATATNCTTCCGCTTCCAGAGAGCG 1249  
1367 AAATCACCGCACTTTGATCTCATGCTGCTGAGTGAATAAGTCCGCTGGTTCGCACTTTG 1426  
1250 TCANCTCGCCCACTACGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1309  
1427 AAATGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1481  
1310 CCAAGCGCCGACCAAGATGAGCAGGGCCCTTGGGAGTTCCGCTCGAGTGTGAAGN 1369  
1482 TCAACCAACATTTGTTTCTTGGTGTGCTGCTGCGGAGAGGACTTCACTTCCAAAGCG 1541  
1370 NNN 1429  
1542 CATCGCCACCGGATTCATTGCGGATCACCCGACCTCTCTCAGGCTCCACCTGCTGATGA 1601  
1430 NNN 1489  
1602 TGACAGGAGCGCATCTCGGATTACTTGGCAGATGTCAACGCTGAAACAGCTCATGGTGT 1661  
1490 NNN 1549  
1662 GCGTCAAGAGTGTGAGCTTCTATCGATGAGTGGCTTAACATCAAGGATCTGCCACT 1721



```

FEATURES
  source          URL:http://fantom.gsc.riken.go.jp/.
  location/Qualifiers
  1..2877
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6J"
    /db_xref="PANTOM,DB:5730513C12"
    /db_xref="MGI:233843"
    /db_xref="taxon:10090"
    /clone="5730513C12"
    /tissue_type="whole body"
    /clone_lib="RIKEN full-length enriched mouse cDNA library"
    /dev_stages="8 days embryo"
  <1..2489
    /note="unnamed protein product; putative
pyruvate decarboxylase [MGI:MGI:97520, GB|NM_008797,
evidence: BLASTN, 9%, match=2877]"
    /codon_start=3
    /protein_id="BAC36939.1"
    /db_xref="GI:26346581"
    /translation="VSSQSLPDLGLROENIRINGCAIQCRVTTEPAPSPQDGRV
EVFRSGEMGIRLNVKAFQAGAVISPHYDLSLLVKVIAHGDHPTAATKMSRALAEPRV
RGVNTNPFQNLVNWQFLAGTVDTQFIDENPELQFLRPAQNRKAKLHLVGHVWV
GPTTPIPVNSPSPFPAVPVPIPPFAGFRDILLREGPEGFARVRNQGILLMDT
TFRAHOSLLATRVTHDLKXIAPYVAHNFKLFMSNMGWGFVDVAMRFLYPCPWR
LOLEETIPNTPFQMLLRGANAVGTNYPDNVFPCFEVAKENGMDVFRVDSLNLYP
NMLQMEAGAGAGVAAISYTGVDADPSRTKYLEVYMLAEELVRACTHILICND
MAGLLKPAACVMTLSLRDLPDLHIHTDTSAGVAAMLACAQAGADVDAVDS
MSGMTSOPSMCALVACTKGTDELVEPLERVDYSEYMGARGLYAADFCTATWKSQ
SDVYNEIRPGQYTNLHFOHSMGLSKFEVKKAYEANKQMLGLDLIKVTPSSKIYGD
LAQMVQNGLSARAEQAELSPFRSVFPLQSTIGIPHGPFPEPRFSKVLKDLPR
EGRPGLPPLNLEKLESLDRHGEVTPEDVLSAAMYPDVFAQKDFATFATGPDLS
LNTLFLQGPXIAEEFEVLEGRKTHIKALAVSDLNAGROVFFELNQLSILVK
DTQAMKENHFPKALDKVQGIAPMPGKVIDIKVAAGDKVAKGQPLCVLSAMKMETV
VTSBMEGTIRKRVHTKDWLTGDDLLILB"
  2851..2856
    polyA_signal
  /note="putative"
  2877
    polyA_site
  /note="putative"

ORIGIN
  Query Match      11.9%; Score 432.2; DB 11; Length 2877;
  Best Local Similarity 53.9%; Pred. No. 8.9e-90;
  Matches 1006; Conservative 0; Mismatches 838; Indels 21; Gaps 5;

QY 1170 CTGGCTGCTGTCGACCTTGAAGGAATTGGCTGTGACCCAGATGAAGATCAAGACCA 1229
DB 2   CGTGTCCGAGGCCAGAGCCCTGCCTGACCTGGGCTGGCGGAGAGAACATCCGCA 61
QY 1230 CGGTGCAGACTGCGAGTCCCGCATCACCAAGAGATCCAAACACCGCTTCCGCCCAGA 1289
DB 62 TGGCTGTGCCATTGCTGTGGGTCAACACCGAGGACCTTGACGAGCTTCCAGCCAGA 121
QY 1290 TACCGGAATATACCGCGTACCGCTCACAGGCGGAGCTGGCTTCTGTGTGACGGTGC 1349
DB 122 CACCGCGCGCATGAGGTTTCCGAGTGGTGGAGGATGGGCATCCGCGCTGGACAACGC 181
QY 1350 AGCTCAGCTCGGTGCGGAATCA---CGCACACTTTTGACTCCATGCTGTGGAATATGAC 1406
DB 182 CTCGTGATTCAGGCGCGTGTATATGCGCCCATATGACTCTCTGTCTGTCTGATGATCAT 241
QY 1407 CTGCGGTGGTTCGCACTTTGAATCTGTGCTGCTGTCACAGCGCGCTTGGCTGAGTT 1466
DB 242 TGCAACAGGCAAAACACCCCGCACAGCTGCCACCAAGATGAGCAGAGCCCTGGCCGAGTT 301
QY 1467 CACGTGTCTGGTGTGCAACCAATGGTTTCTTGGTGGCTGTGGGGAGAGGA 1526
DB 302 CCGTGTCCGAGGTGAAGACCAACATCCCTTCTCGCAAGTGTCTCAACAAACAGCA 361
QY 1527 CTTCACTTCAAGCGCATCGCCACCGGATTCATTCGCCGATCACCGCACCTCTCTTCAGGC 1586
DB 362 GTTCTCGGAGGACAGTGGACACCCAGTTTCATGATGAGAACCTTGAGCTGTTCCAGCT 421

```

```

QY 1587 TCCACTCTGTGATGATGACGAGGAGCGCATCTCTGGATTACTTGGCAGATGTACCGGTGAA 1646
DB 422 TCGGCTGTGACAGAACCGGGGCCGAAAGTTGCTACTATCTCTCGACATGTCTGTGAA 481
QY 1647 CAAGCCTCATGTGTGTCGCTCCAAAGGATGTTGACGCTCTCTATCGAATAAGCTGCCTAACAT 1706
DB 482 TGGCCCTACCACTCCAAATCCCTGTCAATGTGAGCCCGAGTCTGTGGATCTCTGTGTTC 541
QY 1707 CAAGGATC-----TGCCACTGCCAGCGGTTCCGTCAGCCGCTGAAGCAGCTTGGCCC 1760
DB 542 TGTGTGTCCTATGAGCCCACTCCAGCTGGTTTTCAGGACATCTCTTCTCGGAAGAGGGCC 601
QY 1761 AGCCGCGTTTGTCTGTGATCTCCGTGAGCAGCAGCAGCTGAGCAGTTACTGATACCACTT 1820
DB 602 AGAGGCTTTTGGCCGAGCTGTGCGGAATCACCCAGGCGTGTCTGTGTGATGAGACCAACCTT 661
QY 1821 CCGGATGACACACAGTCTTGTGCGACCGAGTCCGCTCATTCGACACTGAAGCCTGC 1880
DB 662 CCGGATGACCCACAGTCACTACTGCGCCACTAGAGTGGCGACACATGATCTCAAAAGAT 721
QY 1881 GGCAGAGGCGGTGCGAAAGCTGACTCTCTGAGCTTTTGTCCGTGAGGCTTGGGGCGGGCC 1940
DB 722 TGGCCCTATGTTGCCCAACAACCTTCAACAGCTCTTCAGCATGGAGAACTGGGAGGCGC 781
QY 1941 GACCTAAGATGTCGATGCGTTTCTCTTTGAGATCCGTGGGACAGCTCGACAGCT 2000
DB 782 CACGTTTCGAGTGTGCGCTTCTCTGTCAGAGTCCGCTTGGCGGCGGCTCCAGAGCT 841
QY 2001 GCGCAGGCGGATGCGCAATGTAAACATTCAGATGCTGCTTCGCGCGCCCAACACCGCTGG 2060
DB 842 CCGGAGCTTATCCCGAACATCCCTTCCAGATGCTACTTGGGGGGCCCATGCTGTGGG 901
QY 2061 ATACACCCGATACCCAGACTCCGCTCTGGCGGCGTTTGTAAAGAACTGCGCAGCTCCGG 2120
DB 902 CTACACCAACTACTCCCTGACCAACGCTGCTTCAAGTCTTGTGAGTGGGCCAAAGAGATGG 961
QY 2121 CTTGACATCTTCCGACCTTCGACGCTTAAACAGCTCTCCAGATGCGTCCAGCAAT 2180
DB 962 TATGAGCTCTTCCGAGTCTTGTGATCTTCCCTTACTTGTCCAAACATGCTGACGGCAT 1021
QY 2181 CGACGAGTCTGGAGACCAACACCGCGGTAGCCGAGTGGCTATGGCTATTCTGGTGA 2240
DB 1022 GGAAGCAG---CAGGCGAGTGTGGGGTGTGGTGGAGGTGCGCATCTCATACCGGGGA 1078
QY 2241 TCTCTGTGATCCAAATGAAAGCTCTACACCTGTGATTAACCTAAAGATGGCAGAGA 2300
DB 1079 CGTGGTGAACCTTAGTTCGACATAAATATCTACTGGAGTACTACATGGGCTTAGCTGAAGA 1138
QY 2301 GATCGTCAAGTCTGGCGCTCAATCTTGGCCATTAAAGATATGGCTGTCTCTTCGCC 2360
DB 1139 ACTGGTGGAGCTGGCACCTCAATCTGTGCATTAAAGACATGGCGGCGCTGCTGAAGCC 1198
QY 2361 AGTGGGTAACAAAGTGTGTCCGCACTGGG---CGTGAATTCGATCTGCCAGTGA 2417
DB 1199 TCCCGCTGACCATGCTGTGTCAGCTCTCCGAGACCGATTCGCCGACCTTCCCACTGA 1258
QY 2418 CGTGACACACCCAGCACACTGCGGGTGGCGAGCTGGCAACCTACTTGTGTCAGCTCAAGC 2477
DB 1259 CATCATACCATGATACATCAGGCGCAGGTGTGGCAGCCATGCTGGCTGTGCACAGC 1318
QY 2478 TGGTGCAGATGCTGTGACGCTGTTCGCACTGTCTGSCACCACTTCCAGCCTATC 2537
DB 1319 AGGGGCTGATGTTGTGGAGCTGGCAGTACACTTGTCTGGGATGACCTTCACAGCCAG 1378
QY 2538 CTTGTCTGCACTTGTGTGCTATTCGCGCACACCGCTGCGATACCGGTTTGGCTCGA 2597
DB 1379 CATGGGGCCCTGTGGTGGCTGTACAAAGGACTCTTTGGACACAGAGGTACCCCTGA 1438
QY 2598 GGTGTTTCTGACCTCGAGCGCTACTGGGAAGCAGTGTGGCGGACTGTACCTGCCATTGA 2657
DB 1439 GGGTGTGTTTGACTACAGTGAAGTACTGGGAAGGGCTCGGGGACTGTACGAGCCTTGA 1498
QY 2658 -----GTCTGGAAACCCAGGCGCCAAACCGGTGCGCTCTACCGCCACGAAATCCAGGCGG 2711

```



QY 1112 TTGAGCACACCGTGAAGAGTCCACCGAGTGGAGCTGTGAAGGCGCAGATCGCT 1171  
 Db 135 TCGAATACAAATTACCGAAGAAATCACCGTATCGATATCGTCTCAATCCAAA 76  
 QY 1172 TGGCTGTGTGCACTTGAAGAAATGGTGTGACCCAGATAGATCAAGACCCAG 1231  
 Db 75 TCGCTGAGGTGCACCTTTTGAACAATTTGGTTGATGCAAGATTAAGATTACAACTCGTG 16  
 QY 1232 GTGCAGCACTGCCAGT 1246  
 Db 15 GTTTCGCCATCCAGT 1

RESULT 6  
 BQ881271  
 LOCUS  
 DEFINITION AGENCOURT 7978339 Lupeki dorsal root ganglion Homo sapiens cDNA  
 clone IMAGE:6185002 5', mRNA sequence.  
 ACCESSION BQ881271  
 VERSION BQ881271.1 GI:22273279  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 AUTHORS NIH-MSC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1AM13575 row: b column: 11  
 High quality sequence stop: 692.

FEATURES  
 Location/Qualifiers  
 1..903  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6185002"  
 /sex="male"  
 /tissue\_type="dorsal root ganglia"  
 /dev\_stage="adult, 36 yr"  
 /lab\_host="DH10B"  
 /clone\_lib="Lupski\_dorsal root ganglion"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
 NotI; Site\_2: SalI; cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCAGCGTCCG-3' and  
 5'-GACTAGTCTAGATCGCAGCGCGCCCT(15)-3'. Size selected >  
 1 kb for average insert length 1.7 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies."

ORIGIN  
 Query Match 6.5%; Score 236; DB 13; Length 903;  
 Best Local Similarity 57.4%; Pred. No. 5.2e-44;  
 Matches 466; Conservative 0; Mismatches 340; Indels 6; Gaps 2;  
 QY 1743 CTTGAAGCAGCTTGCCCGCAGCGCTTTGCTGCTGATCTCCGTGAGCAGGACGCACTGGC 1802  
 Db 18 CTTGCTGGAGAGGGGCTGAGGGCTTTGCTCGACTGTGCGAACCACCGCGGGCTGCT 77  
 QY 1803 AGTTACTGATACCACTTCCCGCGATGCACACAGTCTTTGCTTGGACCCGAGTCCGCTC 1862

Db 78 GCTGATGACACAGCACTTTAGGGAGCGCCCACTACTGCTGGCCACTGCTGTGGCCAC 137  
 QY 1863 ATTGCGCACTGAAGCCTTGGCGCAGAGCGCGTGGCAAAAGCTGACTCCTGAGCTTTTGTCCGT 1922  
 Db 138 CCACGATCTCAAAAGATCGCCCTATGTTGCCCAACAATTTCAGCAAGCTCTTTCAGCAT 197  
 QY 1923 GGAGCGCTGGGGCGCGGACCTAGATGTGGGATGGGTTTCCTCTTTCAGGATCCGTTG 1982  
 Db 198 GGAGAACTGGGGAGGAGCCACGTTTGACGTGCGCATCGGCTTCTGTATGATGCCCTG 257  
 QY 1983 GGACAGGCTCGACGAGTTCGCGGAGGCGATGCCCAATGTAAACATTTCAGATGCTGCTTCG 2042  
 Db 258 GCGCGGCTGCAGAGCTCCGGAGCTCATCCCAACATCCCTTTCCAGATGCTGCTGG 317  
 QY 2043 CGCGCGCAACCGTGGGATACACCCCGTACCCAGACTCCGCTGCGCGGGTTTGTAA 2102  
 Db 318 GGGGGCCAACTCTGTGGCTACACCAACTACCCAGACAAAGTGTCTTCAAGTTCTGTGA 377  
 QY 2103 GGAAGCTGCCAGTCCGGCGTGGACATCTTCCGATCTTTCGACGCGCTTAAACGACGCTC 2162  
 Db 378 AGTGGCCAAAGAGATGGCATGGATGTTCCGTTGTTGACTCCCTCAACTACTTGCC 437  
 QY 2163 CAGATGGCTCAGCAATCGACGAGTCTCTGGAGACCAACACCGCGGTAGCCGAGTGGC 2222  
 Db 438 CAACATGCTGTGGCATGGAGCGG---CAGGAAGTGCAGGCGGTGGAGGCTGC 494  
 QY 2223 TATGCTTTATTTCTGCTGATCTCTCTGATCCAAATGAAAAGCTTACACCCCTGGATTACTA 2282  
 Db 495 CATCTCATACAGGGCGAGCTGGCGCGCCAGCCGACCAAGTACTCACTGCAGTACTA 554  
 QY 2283 CTTAAAGATGGCAGGAGATCGTCAAGTCTGGGCTCATCTTTGGCCATTAAGGATAT 2342  
 Db 555 CATGGCTTGGCCGAAGAGCTGGTGGAGCTGGCACCCACATCTGTGCATCAAGGACAT 614  
 QY 2343 GCTCTGCTGTCTGCGCCAGCTGGGTAACCAAGCTGTGTACCGCACTGCG---CCGTGA 2399  
 Db 615 GCGCGGCTGTCTGAAGCCACAGGCTGCACATGCTGTGCTGCTCTCCGGGACCGTT 674  
 QY 2400 ATTGATCTGCGAGTGCACGTCACACCCAGACATCGGGGTGGCCAGTGGCAACCTA 2459  
 Db 675 CCGCGACCTCCCACTGCACATCCACCCAGACACGTCAGGGGCGAGCGTGGCAGCCAT 734  
 QY 2460 CTTTGTGCGAGCTCAAGCTGGTGGAGATGCTTTCAGCGTCTTCGCGACCACTGCTCTGG 2519  
 Db 735 GTGGCTGTGTCGCGAGCTGGAGCTGATGTGGTGGATGGCAGCTGATTCATGCTCTGG 794  
 QY 2520 CACCACTCCCGAGCCATCCCTGTCTGCCATTG 2551  
 Db 795 GATGACTTCACAGCCAGCATGGGGGCGCTG 826

RESULT 7  
 AQ875053/c  
 LOCUS  
 DEFINITION AQ875053 859 bp DNA linear GSS 08-NOV-1999  
 cerevisiae genomic 5', genomic survey sequence.  
 ACCESSION AQ875053  
 VERSION AQ875053.1 GI:6287297  
 KEYWORDS GSS.  
 SOURCE Saccharomyces cerevisiae (baker's yeast)  
 ORGANISM Saccharomyces cerevisiae  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 REFERENCE  
 AUTHORS 1 (bases 1 to 859)  
 Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,  
 Desreages, S.A., Cheung, K.-H., Sheehan, A., Symoniatidis, D., Jansen, R.,  
 Umansky, L., Heldman, M., Nelson, K., Iwasaki, H., Kanada, D., Iugo, R.,  
 Hager, K., Miller, P., Roeder, G.S. and Snyder, M.  
 TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and  
 Gene Disruption  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Kumar A  
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology

Yale University  
P.O. Box 208103, New Haven, CT 06520-8103, USA  
Tel: 203 432 9949  
Fax: 203 432 6161  
Email: anuj.kumar@yale.edu  
te of mTn-3xHA/lacZ insertion.  
Seq primer: GGCCTTCTTCTTTCGAAGTAC  
Class: transposon-tagged

# FEATURES

source  
1..859  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/strain="Y2278 - S288C background, cir(0) rho(0)"  
/db\_xref="taxon:4932"  
/lab\_host="E. coli"  
/clone\_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"  
/note="Vector: pHS6-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mTn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

# ORIGIN

Query Match 6.2%; Score 223.4; DB 28; Length 859;  
Best Local Similarity 59.9%; Pred. No. 4.5e-41;  
Matches 407; Conservative 0; Mismatches 268; Indels 4; Gaps 2;

QY 827 AAGCTGCTGTGATTACCTCAGCATATTGAAGTCAGATCCTTGGCGATCACTGGAG 886  
DB 801 AATAATCTTGGCAAGAAAGCATATTGAATTTTGGCCGATACACGAAA 742  
QY 887 AAGTTGTACACCTTTATGAACGTGACTGCTACTGCTGAGCGTGTGACCAAAAAGTTTGG 946  
DB 741 CCGTTGTTATCTTTTCGAAAGAGACTGTTTCGCTGCAGAGAAA-CACCAAAAAGTTTGA 683  
QY 947 AATTTGGCGCACACAGCATTTGGATCCAGACTGCTGATCGCATTTGTCGGATGCG 1006  
DB 682 AAGTGGCNCAGCAAGAACTTTACNCGGTGAAGTCGTTGACGCCATTTGACAGATGCG 623  
QY 1007 TAAAGTTTCTGCCCTCCATTTGTTTACAGGGCGCGGAAACCGTGGAAATTTTGGTCGATG 1066  
DB 622 TTAATTTGGCAAGAGTGTGGCTACAGAAATCGCGGTACTGTGTAATTTCTTGGTTGATA 563  
QY 1067 AAAGGGCAACACGCTTTTCATCGAATGAACCCAGTATCCAGGTGAGCACACCGTGA 1126  
DB 562 ACCAAATAGACACTATTTTCAITGAAATTAATCCAAAGATCCAAAGTGAACATACATCA 503  
QY 1127 CTGAAGAAGTCAACGAGGTGGACCTGGTGAAGCGCGAGATGCGCTTGGCTGTGTCGAA 1186  
DB 502 CAGAAGAAATTACCGGTATAGATATTGTGGCGCTCAGATCCAAATTTGGCGAGGTGCT 443  
QY 1187 CTTTGAAGGAATTTGGTCTGACCAAGATGAATCAAGACCCAGTGCAGCACTGCAGT 1246  
DB 442 CTCTACCCAGCTGGGCTATTCCAGACAAAATTAGACTCTGTGCTTTGCCATTCAGT 383  
QY 1247 GCGCATCAACGAGAGATCCAAACACCGCTTCGCGCCAGATACCGGAATTCATCACCG 1306  
DB 382 GCGATTAACCAAGAGACCTGCTAAGAACTTCCAAACAGATACCGGTAGAAATAGAAG 323  
QY 1307 CGTACCGCTACACGGCGAGTGGCTGCTGTGACGCTG---CAGCTCAGCTGGTG 1363  
DB 322 TGTACCGTTCTGCAAGGTGGTAATGTTTATAGCTGGATGTTGTAACGCTATGACGAA 263  
QY 1364 GCGAAATCACCGCACATTTGACTCCATGCTGTGTAATAATGACTGCGGTGGTTCCGACT 1423  
DB 262 CAATAATCTCACTCAATTACGACTCAATGCTGGTCAAAATGCTCATGCTCGGTTCACCT 203  
QY 1424 TTGAATCTGCTGCTGCTGCACAGCGCGGTGGCTGAGTTCACCGTCTCTGTTGTTG 1483  
DB 202 ACGAATCTGTTCTGTAGAAAAATGATTCGTGCTTAATTCGAGTTTCAGAAATAGAGGTGTC 143

QY 1484 CAACCAACATTTGTTTCTT 1502  
DB 142 AGACCAACATTTCCCTTCT 124

# RESULT 8

BF251052  
LOCUS  
DEFINITION  
EST418309 Coccidioides immitis spherule cDNA library Coccidioides immitis cDNA clone CIAAB85 5' sequence, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BF251052.1 GI:16931195  
EST.  
Coccidioides immitis  
Coccidioides immitis  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Onygenales; Mitosporic Onygenales; Coccidioides.

# REFERENCE

1. (bases 1 to 637)  
Gardner M.J. and Kirkland, T.  
Generation of ESTs from Coccidioides immitis spherule cDNA library  
Unpublished (2000)

# AUTHORS

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301 838 3519  
Fax: 301 838 0208  
Email: gardner@igr.org.

# FEATURES

source  
1..637  
/organism="Coccidioides immitis"  
/mol\_type="mRNA"  
/db\_xref="taxon:5501"  
/clone="CIAAB85"  
/dev\_stage="spherule"  
/lab\_host="SOLR"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI"

# ORIGIN

Query Match 6.1%; Score 220.2; DB 10; Length 637;  
Best Local Similarity 60.4%; Pred. No. 2.2e-40;  
Matches 380; Conservative 0; Mismatches 248; Indels 1; Gaps 1;  
QY 685 TACCCCATCTTTTGGAAGCGAGTTGCCGCTGTGGCGGACGCGGTATGCTTTTGTGCT 744  
DB 7 TTCCCATCATCATCAAGCGCGCTTGGCGGCGGTGGCGCGCATGCGTGTGTCGCGC 66  
QY 745 TCACCTGATGAGCTTCGCAATTAGCAACAGAGCATCTCGTGAAGCTGAAGCGGCTTC 804  
DB 67 GACGAGGAGTCCCTCCCGCATCTCTTCGAGCGGCGCACCTCGGAGGCAATCCGCTTC 126  
QY 805 GCGCATGCGCGGTATATGTCGAACGTGCTGTGATTAACCTCAGCATATTGAAGTGCAG 864  
DB 127 GGCACGCGCACCGCTTTCGTCGAGCGCTTCTCGACAGGCGCCCAAGCACATCGAGTCCAG 186  
QY 865 ATCTTGGCGATCACTGAGAAAGTTGTACACCTTTATGAACGTGACTGCTACTGCGAG 924  
DB 187 CTGCTCGCGCAACACCGCGCAACGCTCCTCCACCTTACGAGCGTGTGCTGAGGCTCCAG 246  
QY 925 CGTGTCAACCAAAAGTTGTGAAATTTGCGCAGCACAGCATTTGGATCCAGAACTGCGT 984  
DB 247 CGCCGCCACAGAGGTGCTGAGCTGCGCCCGCCAGGACCTCCCGTCGAGTCCGC 306  
QY 985 GATCGATTTGTGGGATGCAAGTAAGTTTTCGCGCTCCATTTGTTTACGAGGCGCGGGA 1044  
DB 307 GACAGCATCTTCGCGCGAGCGCGTCAAGCTGSCCAAGTCCGTCAGCTACCGCAACGCGCGC 366  
QY 1045 ACCGTGGAATTTCTGTCGATGAAAGGCAACACAGCTTTCATCGAAATGAACCCACGT 1104  
DB 367 ACCGCGGAGTTCTCTGTCGACAGCTCAACGCTACTACTTCTCATCGAGATCAACCCCGC 426

QY 1105 ATCCAGGTTGAGCACACCGTGACGTGAAGAGTCAACCGAGGTGACCTGGTGAAGGCGCAG 1164  
 Db 427 ATCCAGGTGAGCACACCATCACGAGGAGATCACCGGCATCGACATCGTCGCGGCCAG 486  
 QY 1165 ATGCGCTTGGCTGTGTGTGCAACCTTGAAGAAATGGGTCTGACCCCAAGATAGATCAAG 1224  
 Db 487 ATCCAGATCGCGCGCGCGCACCTCTCGAGCAGCTCGGCGCTCACCCAGGACCGCATCTCC 546  
 QY 1225 ACCCAGGTGAGCAGCTGACGTGCGGATCACACCGGAAGAT-CCAAACAACGGCTTCG 1283  
 Db 547 ACCAGAGCTTCGCAATCCAGTGGCGATCACACCGAGATCCCAAGAGGGCTTCCA 606  
 QY 1284 CCCAGATACCGGAATATCACCGGTACC 1312  
 Db 607 GCCGACACCGGTAAAGATCGGAGTCTACC 635

RESULT 9  
 LOCUS CA793824 862 bp mRNA linear EST 04-DEC-2002  
 DEFINITION AGENCOURT 11043532 NICHG\_XGC\_Emb1 Xenopus laevis cDNA clone  
 IMAGE:6865891 5', mRNA sequence.  
 CA793824  
 ACCESSION CA793824.1 GI:26042921  
 VERSION EST  
 KEYWORDS Xenopus laevis (African clawed frog)  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 862)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 plate: L1AM14488 row: p column: 18  
 High quality sequence stop: 649.  
 Location/Qualifiers  
 1..862  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /clone="IMAGE:6865891"  
 /tissue\_type="embryo (stage 10)"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NICHG\_XGC\_Emb1"  
 /notes="vector: pCMV-SPOK6; Site 1: NotI; Site 2: SalI;  
 Cloned unidirectionally. Primer: Oligo dt. Average insert  
 size 1.55 kb. Constructed by Life Technologies. Note: This  
 is a Xenopus Gene Collection (XGC) library."

## FEATURES

source  
 RESULT 10  
 LOCUS AI109467 675 bp mRNA linear EST 19-APR-2001  
 DEFINITION GH08678.5prime GH Drosophila melanogaster head pO2 Drosophila  
 melanogaster cDNA clone GH08678 5prime, mRNA sequence.  
 AI109467  
 ACCESSION AI109467  
 VERSION AI109467.1 GI:3477791  
 KEYWORDS EST.  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 675)  
 REFERENCE 1 Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
 Lewis, S. and Rubin, G.M.  
 BDGP/HMI Drosophila EST Project  
 Unpublished (2001)  
 COMMENT Contact: Stapleton, M.  
 BDGP  
 Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)  
 Plate: 86 row: G column: 6  
 High quality sequence stop: 588.  
 Location/Qualifiers  
 1..675

## ORIGIN

Query Match 6.0%; Score 216.6; DB 14; Length 862;  
 Best Local Similarity 56.7%; Fred. No. 1.8e-39;  
 Matches 421; Conservative 0; Mismatches 39; Indels 3; Gaps 1;  
 QY 841 AACCTTCAGCATATTGAAGTGCAGATCTCTGGCGATCACATCGAGAGTGTACACCTT 900  
 Db 20 AAACCTCCCACTTGAAGTTTCAGATCTCGGTGATAAATATGAAATGTGTACATTG 79  
 QY 901 TATGACGTGACTGCTCACTCAGGCTGTCACCAAAAAGTTGCAAAATGGCCAGCA 960  
 Db 80 TATGAAGAGAGTGTTCGTCCAGAGAGGCGCAGAGGTGGTGAATAGATGCCCGAGCT 139



/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="GH08678"  
/sex="male and female"  
/dev\_stage="adult"  
/lab\_host="DH5 - alpha"  
/clone\_lib="GH drosophila melanogaster head pOT2"  
/notes="Organ: head; Vector: pOT2; Site 1: EcoRI; Site 2:  
XhoI; Sized fractionated cDNAs were directly ligated into  
pOT2. Plasmid cDNA library."

## ORIGIN

Query Match 6.0%; Score 216.4; DB 9; Length 675;  
Best Local Similarity 58.9%; Pred. No. 1.8e-39;  
Matches 392; Conservative 0; Mismatches 271; Indels 3; Gaps 1;  
Qy 1719 ACTGCCACGGTTCCGTGACCGCCCTGAAGACGTTGGCCACCGCGGTGCTCGTGA 1778  
Db 8 ACCACCGAAGGGACTCCGGGAGTCTCTGTGTGCGAAGGTCCGAGGCGCTTCGCCAAGGA 67  
Qy 1779 TCTCCGTGAGCAGGACGACCTGGCAGTTACTGATACCACTTCCGGATGCCACACAGTTC 1838  
Db 68 GTGCGCAACCGTAAGAACTGTACTCATGTGACACCACTTCGCGATGCCACAGTTC 127  
Qy 1839 TTTCTTGGACCGGATGCGCTCATTCGACATGAAGCTTCGCGCAGAGCGCTCGCAA 1898  
Db 128 GCTCTGCGCACCGCGTGGGTTCCACGATCTGTGAAGATCTCCCTTACGTACGCA 187  
Qy 1899 GCTGACTCTGAGCTTTTTCGCGTGGAGGCTGGGGCGGCGACCTACGATGCGCAT 1958  
Db 188 CAAGTTCAACACCTGTATTGCTGTGAGAACTGGGGCGGAGCCACCTTCGACGTGGCGT 247  
Qy 1959 GCGTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTCGCGAGCGCATGCCAA 2018  
Db 248 GCGCTTCTGACAGTGCCTGGGAGCGCTTGGAGAGATCGCAAGCGCATCCGAA 307  
Qy 2019 TGTAAACATTGAGATGCTGCTTGGCGCGCACACCGTGGGATACACCCGTTACCCAGA 2078  
Db 308 CATTCCTTCCAGATGCTGTGGCGGAGCAACGCGCTCGGTACACAGCTATCCGGA 367  
Qy 2079 CTCGCTGCGCGCGTGTGTTAAGGAAGCTGCCAGCTCCGGGCTGGAATCTTCGCGCAT 2138  
Db 368 CAACGTGCTTACAAGTTCTGTGAGCTGGCTGTGCAGACCGGATGGACATCTTCAGGCT 427  
Qy 2139 CTTGACGCGCTTAAAGCTCTCCAGATGCGTCCAGCAATCGACGAGTCTTGAGAC 2198  
Db 428 GTTCGACTCGCTCAACTACCTGCGCAACCTGATCTCCGATGGAAGCGCTCGAAAGGC 487  
Qy 2199 CAACACCGCGTAGCCGAGGTGGCTATGGCTTATTCTGTGATCTCTCTGATCCAAATGA 2258  
Db 488 GGGCGGCTGTGT---GGAGGCTGCCATCTCTATACCGGAGAGCTCAGCGATCCGAGCG 544  
Qy 2259 AAAGCTCTACACCTGGAATTACTACCTAAAGATGGCAGAGAGATGCTGAAGTCTGGCGC 2318  
Db 545 CACCAAGTATGATCTGAATACTACATAACCTTGGCGGATGAGTGTCTCAAGCGCGGCAC 604  
Qy 2319 TCACATCTTGCCCATTAAGGATATGGTGTCTGCTTCCGCGAGCTGCGGTACCAAGCT 2378  
Db 605 CCAAGTCTCTGATTAAGGATATGGTGTCTGCTCTCAAGCCGATCCGCGAGGCTTCT 664  
Qy 2379 GGTGAC 2384  
Db 665 GATCAC 670

RESULT 11  
BI103175  
LOCUS 744 bp mRNA linear EST 26-JUN-2001  
DEFINITION 602889415f1 NCI\_CGAP\_Kid14 Mus musculus cDNA clone IMAGE:5044779  
5', mRNA sequence.  
ACCESSION BI103175  
VERSION BI103175.1 GI:14554068

## KEYWORDS

Mus musculus (house mouse)

## SOURCE

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 744)

## REFERENCE

NIH-MGC <http://mgc.nci.nih.gov/>.

## AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

## TITLE

Unpublished (1999)

## JOURNAL

Contact: Robert Strausberg, Ph.D.

## COMMENT

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM1122 row: m column: 04  
High quality sequence stop: 737.

## FEATURES

Location/Qualifiers

1..744

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:5044779"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NCI CGAP Kid14"

/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.75 kb. Constructed by Life

Technologies. Note: this is a NCI\_CGAP Library."

## ORIGIN

Query Match 5.9%; Score 214.8; DB 12; Length 744;  
Best Local Similarity 58.6%; Pred. No. 4.4e-39;  
Matches 430; Conservative 0; Mismatches 257; Indels 7; Gaps 3;  
Qy 1765 GCGTTTGTCTGTGATCTCCGTGAGCAGGACGACATCGGAGTCTGTTGATGGACACAACTTCGG 74  
Db 15 GCGTTTGTCTGTGATCTCCGTGAGCAGGACGACATCGGAGTCTGTTGATGGACACAACTTCGG 74  
Qy 1825 GATGCACACCACTTTTGTGTCGACCGAGTCCGCTCATTCGCACTGAAGCTTCGCGCA 1884  
Db 75 GATGCCACCACTTACTACTGCGCCTAGAGTGGCACAATGATCTCAA-AAAGATGGC 133  
Qy 1885 GAGCGCGTCGCAAGCTGACTCTGAGCTTTTGTTCGTGAGGCTTCGGGCGGCGGAC 1944  
Db 134 CCTATGTGTCGCCCAACTTCAACAAGCTCTTCAGCATGGAGAACTGGGAGGCGCCACG 193  
Qy 1945 TACGATGTGGCGATCGCTTCTTTCGATCCGTCGCGACAGCTCGAGCTCGCGC 2004  
Db 194 TCGACGTGTGCATCGCTTCTGTACAGTCCCTTGGCGGCGCTCCAGAGCTCGG 253  
Qy 2005 GAGCGGATCCGGAATGTAAACATTCAGATGCTGCTTCGCGCCGCAACACCGTGGGATAC 2064  
Db 254 GAGCTTATCCGCAACATCCCGTTCAGATGCTACTGAGGGGGGCGCAATGCTGTGGGTAC 313  
Qy 2065 ACCCGTACCCAGACTCGCTCTGCGGCGCTTTGTAAGGAAGCTGCCAGCTCCGCGGTG 2124  
Db 314 ACCAACTACCTTGACACAGTGTCTTCAAGTCTGTGAGGTGGCCAAAGAGATGGTATG 373  
Qy 2125 GACATCTTCCGCATCTTCGACGCGCTTAAACAGCTCTCCAGATGCGTCCAGCAATCGAC 2184  
Db 374 GAGGTCTTCCGAGTCTTTGACTCTCCTCACTTTCGCAACATGCTGCTGGGATGGAA 433  
Qy 2185 CGAGTCTTGGAGACCAACACCGGTAGCGGTGCTTATGCTTATTCGTGTGATCTC 2244  
Db 434 GCAG---CAGCGAGTGTGGGGGTGTGTGGAGGCTGCTCATACACGCGGAGACGTG 490  
Qy 2245 TCTGATCCAAATGAAAGCTCTACACCTCGGATCTACACCTGATCTACTACCTAAAGATGGCAGAGATC 2304

Db 491 GCTGACCTAGTCCGCACTAAATCTACTGAGTACTACATGGGCTTAGCTGAAGAACTG 550

Qy 2305 GTCAAGTCTGGCGCTCACAATCTTGGCCATTAAGGATATGGCTGTCTGTCCCGCAGCT 2364

Db 551 GTGCGAGCTGGCACTCACAATCTTGGCCATTAAGGATATGGCTGTCTGTCCCGCAGCT 610

Qy 2365 GCGGTAAACCAAGCTGTCACCGCACTGG---CCGTGAATTCGATCTGCGCAGTGCACG 2421

Db 611 GCTGCAACCAATGCTGGTGGTCCCTCCCGGACCGATTCGCCGACCTCCCACTGCACATC 670

Qy 2422 CACACCCACCACTGCGGGGTGGCCAGCTGGCAACCTACTTTGTGTCAGCTCAAGCTGGT 2481

Db 671 CATACCCATGATACATCAGGGGCGAGGTGGCGACCATGCTGGCGCTGTGCACAGCAGGG 730

Qy 2482 GCAGATGCTGTGGA 2495

Db 731 GCTGATGTCGTGGA 744

## RESULT 12

CK000115 800 bp mRNA linear EST 26-NOV-2003

LOCUS AGENCOURT\_16368972 NIH MGC 220 Homo sapiens cDNA clone

DEFINITION IMAGE:30707731 5', mRNA sequence.

CK000115

CK000115.1 GI:38526149

## ACCESSION

EST.

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 800)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: James Martin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/ILLNL at:

<http://image.llnl.gov>

Plate: NDAM1072 row: m column: 20

High quality sequence start: 12

High quality sequence stop: 654.

Location/Qualifiers

1..800

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30707731"

/lab\_host="DH10B TconA"

/clone\_lib="NIH\_MGC\_220"

/note="Organ: mixed; Vector: pYX-Asc; Site: 1: EcoRI;

Site 2: NotI; Library is oligo-dT primed and directionally

cloned. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated with

EcoR I adaptor, digested with Not I and then cloned

directionally into pYX-Asc vector. Average insert size

0.5-1kb. Adaptors 5' (AATTCGACAGG)3' and 5'd

(CTCTGTGCG)3'.

3' Linker sequence - CGGCGCGTGAGAGCC T18.

Sequencing primers 3'end: T3 promoter primer 5'd

(ATTACCTCTACTAAGGA)3'.

5' End: T7 promoter primer 5'd

(TAATACGCTACTATAGG)3'.

Library was constructed in the

laboratory of M. Bento Soares.

Average insert size 3-4kb

Note: this is a NIH\_MGC Library."

## FEATURES

source

## ORIGIN

Query Match 5.9%; Score 213.2; DB 14; Length 800;

Best Local Similarity 59.1%; Pred. No. 1.1e-38;

Matches 401; Conservative 0; Mismatches 273; Indels 4; Gaps 2;

Qy 841 AACCTCTAGCATATTGAAGTGCAG-ATCCTTGGCGATCACACTGGAGAAAGTTGTACACCT 899

Db 35 AGCCACCGGCATCTGAGTGCAGAACTTGGGGGACCAAGTATGGAAACATCTTGCACCT 94

Qy 900 TTATGAACGTGACTGCTCTCACTGCGAGCTGTCACAAAGTTGTGAAATTCGAGCCAGC 959

Db 95 GTACGCGAGAGAGTCTCCATCCAGCGCGCGCACAGAAAGTGTTCGAGATTGCCCGCG 154

Qy 960 ACAGCATTTGGATCCAGAACTGCGGTGATCGCATTTGTGCGGATGACGTAAGAGTTCTGCG 1019

Db 155 CGCCCACTGGACCGCGAGCTTCGAGCTTCAGCGGCTCACAGCGACTCTGTGAAACTGCGTAA 214

Qy 1020 CTCCTATTGTTTACCAAGGCGCGGGAACCGTGGAAATTTTGTTCGATGAAAGGCAACCA 1079

Db 215 ACAGGTGGGGTAGGAGAACGCGAGGCAACCGTGGAGTTCTCTGGTGGACAGGACCGCAAGCA 274

Qy 1080 CGTCTTCATCGAAATGAACCCACGATATCCAGTTTGAGCAGACACCGTACTGAGAGTCCAC 1139

Db 275 CTACTTCTATCGAGGTCAACTCCCGCTCGAGTGGAGCACACGGTCCAGAGGAGATCAC 334

Qy 1140 CGAGGTGACCTGTGTGAAGGCGCAGATCGGCTTGGCTGCTGTCACACTTGAAGGAAAT 1199

Db 335 CGACGTAGACCTGTGTCATGCTCAGATCCACGTGGCTGAGGCGAGGAGCTTACCGGACT 394

Qy 1200 GGGTCTGACCCCAAGATAGATCAAGACCCAGTCAGTGCAGACTGCGAGTCCGCATCACCAC 1259

Db 395 GGGCTCTGGCGGAGGAGAAATCCGATCAACGGGTGTGCCATCCAGTCCCGGTCAACAC 454

Qy 1260 GGAAGATCCAAACAAACGGCTTCCGCCAGATACCGGAACTATCACCGGTACCGCTCAC 1319

Db 455 CGAGAACCCCGCGCGAGCTTCCAGCGGACACCGCGCGCATTTGAGTGTTCGGAGCGG 514

Qy 1320 AGCGGAGCTGCGGTTCGCTCT---TGACGCTGACGCTGCGTGGTTCGAACTTTGAAACTGCTGT 1376

Db 515 AGAGGCGCATGGGCATCCGCTCGGTAATGCTTCCGCTTCCAGGAGCGCTCATCTCGCC 574

Qy 1377 ACATTTGACTCCATGCTGCTGAAATGACCTGCGTGGTTCGAACTTTGAAACTGCTGT 1436

Db 575 CCATACGACTCTCTGCTGCTCAAGTCAATGCTCCACGCAAGACACCCACCGCGCGC 634

Qy 1437 TGCTGTCACACGCGCGCTTGGCTGAGTTCACCGTGTCTGTGTCACCAACCAATTGG 1496

Db 635 CACCAAGATGAGCAGGCGCGCTTGGGAGTTCGCGGCTTCGAGGTGTGAAGACCAACATCGC 694

Qy 1497 TTTCTTGGCTGGCTTGCCT 1514

Db 695 CTTCTCTGCAAGATGTGCT 712

## RESULT 13

BI618593

LOCUS

DEFINITION

BI618593.1 GI:15514118

EST.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 640)

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B.,

Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E.,

located on: 2R 46B14-46C1; 08/19/2001, mRNA sequence.

BI618593

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BI618593 640 bp mRNA linear EST 07-SBP-2001

RF49459.5prime RH Drosophila melanogaster normalized Head pFlc-1

Drosophila melanogaster cDNA clone RH49459 5 similar to

BCDNA:GH06348: FBAN0001516 Go: lpyruvate carboxylase (GO:0004736)]

located on: 2R 46B14-46C1; 08/19/2001, mRNA sequence.

BI618593

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C.J., Nuno, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin, G.M.

BDGP/HMI RH Drosophila EST Project

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)

hit genomic AB003831: arm:2R [4876906..5137815]

estimated-cyto:46B13-46E4: 08/19/2001

Plate: RH.494 row: E column: 11

High quality sequence stop: 543.

FEATURES  
source

1. .640  
Location/Qualifiers  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="RH49459"  
/sex="male and female"  
/dev\_stage="Adult"  
/lab\_host="DH5-alpha Tona"  
/clone\_lib="RH Drosophila melanogaster normalized head pFic-1"

/notes="Organ: head; Vector: pFic1; Site 1: XhoI; Site 2: BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

ORIGIN

Query Match 5.8%; Score 211; DB 12; Length 640;  
Best Local Similarity 59.4%; Pred. No. 3.2e-38;  
Matches 377; Conservative 0; Mismatches 255; Indels 3; Gaps 1;  
Qy 790 GCTGAACGGCTTCGCGATGCGGGTATATGTCGACGCTGCTGATTACCTCAG 849  
Db 1 GCCAGGCCCGCTTGGCAATGGTGCATGTTCAITGAGAGTTCAATCAGGACCGGT 60  
Qy 850 CATATTGAAGTCAGATTCCTTGGCGATCACTATGGAGAAGTTGTACACCTTTATGAAC 909  
Db 61 CACATTGAGTTCACTGCTGGGAGACAGAGCGCGGAACTAGTGCATCTGTACGAAC 120  
Qy 910 GACTGCTCACTGAGCGTCTCCACCAAAAGTTGTCBAATTCGCCAGCAGCATTTG 969  
Db 121 GACTGCTCCGTGAGCGTCCGACAGAGTGGTGGAGATCGCCCGAGCGGACGCGCT 180  
Qy 970 GATCCAGAACTGGTGATGTCATTTGTGCGGATGCGAGTAAAGTTCTGCCGCTCCATT 1029  
Db 181 CCCATCGAAATCCGGGACAAGATGACGGAGGACGAGTGGCTTGGCCCGCATGTGG 240  
Qy 1030 TACAGGCGCGGGAACCGTGGATTTCTTGGTGCATGAAAGGGAACCAACGCTTTCAT 1089  
Db 241 TATGAAACCGCGAACCGTGGAGTTCCTTTGGAGAGTCCGCGCAACTTCTACTTCAT 300  
Qy 1090 GAAATGAACCCACGATATCCAGGTTGAGCACACCGTCACTGAAGAAGTCAACGAGTGG 1149  
Db 301 GAAGTGAACGCCCTCTACAAGTGGAGCAACCGTCAACCGAGAGATAACAGGATCGAT 360  
Qy 1150 CTGTGAAGGGCGAGATCGCGTGTGCTGTGGTGCACCTTTGAAGGAATGGGTCTGACC 1209  
Db 361 TTGTGTCAGTGCAGAAATCCGCGTGGCGGAGGATGACCCCTACCCGAAATGGGCTAC 420  
Qy 1210 CAAGATAAGATCAAGACCAACCGGTGAGCACTGTCAGTGGCGCATCACCGGAAGATCCA 1269  
Db 421 CAGACAGATCGTGGCGGTGGCTATGCCATTCACTGTCGCTGACCCAGGATCCG 480  
Qy 1270 AACACGGCTTCGCCCCAGATACCGGAATATACACCGCGTACCGCTCACCGCGGAGCT 1329  
Db 481 GCCAACGATTTTACAGCCCAACACCGGTCGCTTGGAGTCTTCCGATCTGGCGAGGTATG 540

Qy 1330 GGGGTTCGCTTTCACGGTG--CAGCTCAGCTGGTGGGGAATCACCACACATTTGAC 1386  
Db 541 GGCATAGTTGACAGTGGTCCGCTAGCGCGAGGACATCATTTGGCCATACAGAC 600  
Qy 1387 TCCATGCTGCTGAAATGACCTGCCGTGCTGCGA 1421  
Db 601 TCCCTGCTGCTCAAGGTGATCTCGCACGCCAGCGA 635

RESULT 14

BG910468

LOCUS

DEFINITION

602806142F1 NCI\_CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4938203

5', mRNA sequence.

ACCESSION

BG910468

VERSION

BG910468.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 715)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: David N. Louis, M.D.

CNSA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLNMI0873 row: d column: 12

High quality sequence stop: 715.

FEATURES

Location/Qualifiers

1. .715

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="IMAGE:4938203"

/tissue\_type="anaplastic oligodendroglioma with 1p/19q

loss"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NCI\_CGAP\_Brn67"

/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 2.3 Kb. Constructed by Life

Technologies. Note: this is a NCI\_CGAP Library."

ORIGIN

Query Match 5.7%; Score 208.2; DB 12; Length 715;  
Best Local Similarity 56.8%; Pred. No. 1.5e-37;  
Matches 403; Conservative 0; Mismatches 303; Indels 3; Gaps 1;  
Qy 872 GCGATCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTCAGCGTGC 931  
Db 1 GGGACCAATGAGAAATCTGTCACCTGTACGAGCGAGACTGTCTCCATCAGCGCGGC 60  
Qy 932 ACCAAGAGTTGCGAATTCGCCAGACACAGCATTTGGATCCAGAACTCGTGTATCGCA 991  
Db 61 ACCAAGAGTTGCGAATTCGCCAGACACAGCATTTGGATCCAGAACTCGTGTATCGCG 120  
Qy 992 TTTGTGCGAGTGCAGTAAAGTTCTGCCGCTCCATTTGGTTACAGGGCGCGGAAACCGTGG 1051  
Db 121 TCACCAGCGACTCTGTGAACCTCGCTAAACAGAGTGGGCTACGAGAACGACGACCGTGG 180  
Qy 1052 AATCTTTGGTGCAGAAAGCGCAACCGCTCTTCATCGAATAAGAACCCGATCCAGG 1111  
Db 181 AGTTCTTGGTGCAGCGACGCGCAAGCACACTTCTCATCGAGGTCACTCCCGCTTCAGG 240  
Qy 1112 TTGAGCACACCGTGAAGAGTCAACCGAGGTGAGACCTGGTGGAGCGCGAGATGGCT 1171

Db 241 TGGAGCACACGGTCACAGAGGAGATCACCGAGTAGACCTGGTCCATGCTCAGATCCACG 300  
 QY 1172 TGGCTGCTGGTGCACACCTTGAAGGAATTGGTCTGACCCAGATAAGATCAAGACCCACG 1231  
 Db 301 TGGCTGAGGGCAGGAGCTTACCGACCTTGGCCCTTGGGCGAGAGAAATCCGATCAACG 360  
 QY 1232 GTGCAGCACTCAGTGGCGGATCAACCGAAGATCCAAACACAGGGCTTGGCCAGATA 1291  
 Db 361 GTGTGGCCATCAGTGGCGGGTCAACACCGAGGACCCCGCGCAGCTTCCAGCCGACA 420  
 QY 1292 CCGGAATATCACCGGTACCGCTCACCGAGGGAGCTGGGTTGCTTTCACGGTGCAG 1351  
 Db 421 CCGGCCGCAATGAGGTGTCGGAGCGGAGGGCATGGGCATCCGCTCGATTAATGCTT 480  
 QY 1352 CTCAGCTC---GGTGGCGAAATCACCGCACACTTTTGACTCCATGCTGCTGAAATGACCT 1408  
 Db 481 CGGCTTCCAAAGGAGCGCTCATCTGCCCCACTAGACTCCCTGCTGCTCAAGTCAATTG 540  
 QY 1409 GCCGTGGTTCGACTTTGAAACTGCTGTGTCTGTGCACAGCGCGCTTGGCTGAGTTCA 1468  
 Db 541 CCCAGCGAAGAGCAACCCACGGCGCCACCAAGATGAGCAGGGCCCTTGGCGAGTTCC 600  
 QY 1469 CCGTGTCTGGTGTGCAACCAACTTGGTTTCTTGGTGGCTGCTGCGGAGAGGACT 1528  
 Db 601 GGTCCGAGGTGTGAGACCACTGCTGCTTCTGCGAATGTGCTCAACACCGAGCT 660  
 QY 1529 TCACCTTCAAGCGATCCGACCGGATTCATGCGGATCACCCGCACT 1577  
 Db 661 TCCTGGCAGGCACTGTGTGACACCACTTTCATCGACGCGAGACCCAGAGCT 709

RESULT 15  
 BJ063722 641 bp mRNA linear EST 29-SEP-2003  
 LOCUS BJ063722 NIBB Mochii normalized Xenopus tailbud library Xenopus  
 DEFINITION laevis cDNA clone XJ077c10 5', mRNA sequence.  
 ACCESSION BJ063722  
 VERSION BJ063722.1 GI:17470882

KEYWORDS EST.  
 SOURCE Xenopus laevis (African clawed frog)  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 Xenopodinae; Xenopus.  
 REFERENCE 1 (bases 1 to 641)  
 AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y

TITLE Expressed genes in X. laevis embryo  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Tadasu Shin-i  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tsunigenes.nig.ac.jp  
 The information of this clone is available through the following URL.  
 http://xenopus.nibb.ac.jp.

FEATURES  
 source  
 1. 641  
 Location/Qualifiers  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /clone="XJ077c10"  
 /tissue\_type="whole embryo"  
 /dev\_stage="stage 25"  
 /clone\_lib="NIBB Mochii normalized Xenopus tailbud library"

ORIGIN

Query Match 5.6%; Score 204.4; DB 12; Length 641;

Best Local Similarity 58.9%; Pred. No. 1.1e-36;  
 Matches 371; Conservative 0; Mismatches 256; Indels 3; Gaps 1;  
 QY 841 AACCTCAGCATATTGAAAGTCAGATCTTGGCGATCACTGGAGAGATTGTACACCTT 900  
 Db 5 AACCTCGCCACATTGAAGTTCAGATACCTCGGTGATAAATATGGAATGTGTGATATTG 64  
 QY 901 TATGAACGTGACTGCTCACTGAGCGGTGCTCAACCAAAAAGTTTGTGAAAATGCGCCAGCA 960  
 Db 65 TATGAAGAGAGACTGTTCCGTCCAGAGAAGGCACCAAGAGTGGTTGAAATAGCCCCAGCT 124  
 QY 961 CAGCATTTGGATCCAGAACTGGTGTATGCGATTTGTGGGATGCACTAAAGTTCTGCCGC 1020  
 Db 125 GCACAGCTGACCTTCAGCTCAGACACAGACTGACCCATGACTCTGTATAACTGGCAAAA 184  
 QY 1021 TCCATTGGTTTACAGGGCGCGGAAACCGTGGAAATTTCTTGGTTCGATGAAAAGGGCAACCAC 1080  
 Db 185 CAGGTGGGTTATGAGAAATGCGCGCACTGTGAGATTCTTGGTGGATAAACATGCGCAAGCAT 244  
 QY 1081 GTCTCATCGAAATGAACCCAGTATCCAGGTTGAGCACACCGTGTGAGAAAGTCAACC 1140  
 Db 245 TACTTTATAGAAAGTCAACTCCAGACTCCAAGTTGAGCATACAGTAACCGAAGAAATTACA 304  
 QY 1141 GAGGTGGACCTGCTGAAGCGCGAGATGGCTTGGCTGTGTTGCAACCTTGAAGGAATTG 1200  
 Db 305 GATGTTGACCTAGTCCATGCCAGATCCATGTTGAGAGGGAAGAAAGCTTCTCTGACCTT 364  
 QY 1201 GGTCTGACCCAAAGATAAGATCAAGACCCAGCGTGGCAGCACTGAGTGGCGCATCAACG 1260  
 Db 365 GGCCTCAAGCAAGAGAACATCCGTATCAATGGTTGTGCCATTTCAGTGCAGAGTCAACA 424  
 QY 1261 GAAGATCCAAACAACAGCGGCTTCGCCCCAGATACCGGAACCTATCACCGCGTACCGCTCACCA 1320  
 Db 425 GAAGACCTTTCAGTGGGTTTCAGCTGATCTGTCGGATTGAGGTTTTCGGAGTGA 484  
 QY 1321 GCGGAGCTGGCGTTTCGTCTTGACGGTG---CAGCTCAGCTCGGTGGCGAAATCAACGCA 1377  
 Db 485 GAAGGAATGGGAATCCGACTAGACAGTGTTCAGCTTTCAAGGAGCTGTCAATTTCCCA 544  
 QY 1378 CACTTTGACTCCATGCTGTGAAATGACCTGCGGTTCGAGCTTTGAACTGCTGTT 1437  
 Db 545 CATTATGACTCCCTTCTGTTCAAGGTTATTGACACCGGAAAGACCATCAAGTTGCCGCC 604  
 QY 1438 GCTCGTGCACAGCGCGCTTGGCTGAGTTTC 1467  
 Db 605 ACAAAAATGAGCAGAGACTTTGGCAGAGTTC 634

Search completed: March 24, 2004, 02:37:29  
 Job time : 8678 secs